THE QUARTERLY JOURNAL OF MATHEMATICS

OXFORD SERIES

Volume 17 *No.* 68 *December* 1946

CONTENTS

OXFORD
AT THE CLARENDON PRESS
1946

Price 7s. 6d. net

PRINTED IN GREAT BRITAIN BY CHARLES BATEY AT THE OXFORD UNIVERSITY PRESS

LOMMEL'S FUNCTIONS OF SMALL ARGUMENT

By F. B. PIDDUCK (*Oxford*)

[Received 10 May 1945]

LOMMEL* found two particular integrals of the differential equation

$$\left(z^2\frac{d^2}{dz^2}+z\frac{d}{dz}+z^2-n^2\right)y = z^{m+1},$$

which is Bessel's equation when the right-hand side is zero, in the form

$$S_{m,n}(z) = z^{m-1}\left[1-\frac{(m-1)^2-n^2}{z^2}+\frac{\{(m-1)^2-n^2\}\{(m-3)^2-n^2\}}{z^4}-\ldots\right],$$

$$s_{m,n}(z) = \frac{z^{m+1}}{(m+1)^2-n^2}-\frac{z^{m+3}}{\{(m+1)^2-n^2\}\{(m+3)^2-n^2\}}-\ldots.$$

The first suffers from the disadvantage that the series diverges unless it terminates, while the second is liable to a number of causes of failure. But if $S_{m,n}(z)$ is defined by the equation

$$S_{m,n}(z) = s_{m,n}(z)+2^{m-1}\Gamma(\tfrac{1}{2}m-\tfrac{1}{2}n+\tfrac{1}{2})\Gamma(\tfrac{1}{2}m+\tfrac{1}{2}n+\tfrac{1}{2})\times$$
$$\times[\sin(\tfrac{1}{2}m\pi-\tfrac{1}{2}n\pi)J_n(z)-\cos(\tfrac{1}{2}m\pi-\tfrac{1}{2}n\pi)Y_n(z)], \quad (1)$$

there is no difficulty in the case of failure, and $S_{m,n}(z)$ can always be defined.

Lommel's functions owe a great deal of their importance to a formula which I shall write

$$\int\limits_0^z z^m J_n(z)\,dz = I_{m,n}(z)-I_{m,n}(0), \quad (2)$$

where

$$I_{m,n}(z) = (m+n-1)zJ_n(z)S_{m-1,n-1}(z)-zJ_{n-1}(z)S_{m,n}(z). \quad (3)$$

The integrals as given by Lommel are indefinite, but the lower limit 0 is often wanted, and that requires special investigation. This is the object of this paper. We suppose that m and n are positive integers (including zero).

* E. Lommel, *Math. Ann.* 9 (1876), 425. The work is reproduced with improvements and additions by G. N. Watson, *Bessel Functions*, 2nd ed., pp. 345–52. These pages contain all the formulae we need, and will be used without further reference.

The sign $\approx$ referring to approximate values when z is small, we have

$$\left.\begin{aligned} J_n(z) \approx z^n/n!\,2^n, \qquad J_{-n}(z) \approx (-)^n z^n/n!\,2^n, \\ Y_0(z) \approx 2\log(\tfrac{1}{2}z)/\pi, \qquad Y_n(z) \approx -(n-1)!\,2^n/\pi z^n \quad (n \geqslant 1), \\ Y_{-n}(z) \approx -(-)^n(n-1)!\,2^n/\pi z^n \quad (n \geqslant 1). \end{aligned}\right\} \tag{4}$$

If $m-n$ is not an odd negative integer,

$$s_{m,n}(z) \approx z^{m+1}/(m+n+1)(m-n+1). \tag{5}$$

Substituting these values we find approximate values of $I_{m,n}(z)$ when $m \neq 1$ and $n = 0$, and when $m-n$ is not an odd negative integer and $n > 1$. These values have a finite limit when $z \to 0$, and thus we find

$$\left.\begin{aligned} I_{m,0}(0) &= -2^{m-1}(m-1)\cos\tfrac{1}{2}m\pi\Gamma(\tfrac{1}{2}m+\tfrac{1}{2})\Gamma(\tfrac{1}{2}m-\tfrac{1}{2})/\pi \\ &= -2^m\Gamma(\tfrac{1}{2}m+\tfrac{1}{2})/\Gamma(-\tfrac{1}{2}m+\tfrac{1}{2}) \quad (m \neq 1), \\ I_{m,n}(0) &= -2^m\cos(\tfrac{1}{2}m\pi-\tfrac{1}{2}n\pi)\Gamma(\tfrac{1}{2}m-\tfrac{1}{2}n+\tfrac{1}{2})\Gamma(\tfrac{1}{2}m+\tfrac{1}{2}n+\tfrac{1}{2})/\pi \\ &= -2^m\Gamma(\tfrac{1}{2}m+\tfrac{1}{2}n+\tfrac{1}{2})/\Gamma(-\tfrac{1}{2}m+\tfrac{1}{2}n+\tfrac{1}{2}) \\ &\qquad (m-n \text{ not an odd negative integer, } n > 1). \end{aligned}\right\} \tag{6}$$

The failing cases are considered conveniently under four heads: (*a*) $m = 1$, $n = 0$; (*b*) $m = 0$, $n = 1$; (*c*) $m = n-1$ in general; (*d*) $m = n-2p-1$, where p is a positive integer.

(*a*) $I_{1,0}(z) = -zJ_{-1}(z)S_{1,0}(z)$. This is only an *apparent* failing case, since the coefficient of $S_{m-1,n-1}(z)$ vanishes. Since $s_{1,0}(z) \approx z^2/4$, we find $S_{1,0}(z) \approx 1$, and thus

$$I_{1,0}(0) = 0. \tag{7}$$

(*b*) $I_{0,1}(z) = -zJ_0(z)S_{0,1}(z)$, and

$$S_{0,1}(z) \approx \tfrac{1}{4}z\Gamma(1)\{2\log(\tfrac{1}{2}z)-\psi(2)-\psi(1)\}/\Gamma(2)-\tfrac{1}{2}\pi\Gamma(1)[-2/\pi z] \approx z^{-1}.$$

Hence

$$I_{0,1}(0) = -1. \tag{8}$$

(*c*) $I_{n-1,n}(z) = (2n-1)zJ_n(z)S_{n-2,n-1}(z)-zJ_{n-1}(z)S_{n-1,n}(z)$. The case $n = 1$ has just been considered. We need not consider the case $n = 0$, since (2) shows that the integral does not converge. Otherwise we take the obviously largest terms in Watson's expressions for $S_{\mu,\nu}(z)$ in the failing case $\mu = \nu-1$, and find

$$\begin{aligned} S_{n-2,n-1}(z) \approx \tfrac{1}{4}z^{n-1}\Gamma(n-1)\{2\log(\tfrac{1}{2}z)-\psi(n)-\psi(1)\}/\Gamma(n)- \\ -2^{n-3}\pi\Gamma(n-1)[-(n-2)!\,2^{n-1}/\pi z^{n-1}], \\ S_{n-1,n}(z) \approx \tfrac{1}{4}z^n\Gamma(n)\{2\log(\tfrac{1}{2}z)-\psi(n+1)-\psi(1)\}/\Gamma(n+1)- \\ -2^{n-2}\pi\Gamma(n)[-(n-1)!\,2^n/\pi z^n]. \end{aligned}$$

Hence

$$I_{n-1,n}(0) = -2^{n-1}\Gamma(n) \quad (n \geqslant 1). \tag{9}$$

(d) $I_{n-2p-1,n}(z) = (2n-2p-1)zJ_n(z)S_{n-2p-2,n-1}(z)-zJ_{n-1}(z)S_{n-2p-1,n}(z)$, and

$$S_{n-2p-2,n-1}(z) \approx z^{n-2p-1}/4(-p)_1(n-1-p)_1- \\ -2^{n-3}\pi\Gamma(n-1)[-(n-2)!\,2^{n-1}/\pi z^{n-1}],$$

$$S_{n-2p-1,n}(z) \approx z^{n-2p}/4(-p)_1(n-p)_1-2^{n-2}\pi\Gamma(n)[-(n-1)!\,2^n/\pi z^n],$$

in each of which the second term is of larger order than the first when z is small. Thus we have the same result as in case (c),

$$I_{n-2p-1,n}(0) = -2^{n-1}\Gamma(n) \quad (p > 0). \tag{10}$$

The cases partly overlap or contain each other, but have been taken separately since the mode of proof is not always the same.

To illustrate the use of the formulae, consider the function

$$E_n^s(x) = \int\limits_0^\pi J_0(xy)\exp(iny)y^s\,dy, \tag{11}$$

where n is an integer, s a positive integer, and x is real. For moderate values of x we can expand $J_0(xy)$ in powers of xy and get an ascending power series for $E_n^s(x)$, with coefficients that can be calculated numerically, though only by a professional computer. If we expand the exponential in powers of y, and change the variable from y to xy, we have quite generally

$$\frac{E_n^s(x)}{\pi^{s+1}} = \sum_{m=0}^{\infty}\frac{(in\pi)^m}{m!(\pi x)^{m+s+1}}\times \\ \times[(m+s-1)zJ_0(z)S_{m+s-1,-1}(z)+zJ_1(z)S_{m+s,0}(z)]_0^{\pi x}. \tag{12}$$

If x is large, we use the known asymptotic expansions of $S_{m+s-1,-1}(z)$ and $S_{m+s,0}(z)$. These give series which can be summed with respect to m, and a little reduction gives the following approximate values when x is large:

$$E_1^0(x) \approx (x^2-1)^{-\frac{1}{2}}+ \\ +\pi J_0(\pi x)\left[\frac{1-\pi i}{(\pi x)^2}+\frac{-3+2\pi^2+(3\pi-\pi^3)i}{(\pi x)^4}+\right. \\ +\frac{45-24\pi^2+3\pi^4+(-45\pi+9\pi^3-\pi^5)i}{(\pi x)^6}+ \\ \left.+\frac{-1575+810\pi^2-78\pi^4+4\pi^6+(1575\pi-285\pi^3+18\pi^5-\pi^7)i}{(\pi x)^8}+\cdots\right]+ \\ +\pi J_1(\pi x)\left[\frac{-1}{\pi x}+\frac{1-\pi^2+(-\pi)i}{(\pi x)^3}+\frac{-9+5\pi^2-\pi^4+(9\pi-2\pi^3)i}{(\pi x)^5}+\right. \\ \left.+\frac{225-117\pi^2+12\pi^4-\pi^6+(-225\pi+42\pi^3-3\pi^5)i}{(\pi x)^7}+\cdots\right],$$

$$E_1^1(x) \approx -(x^2-1)^{-\frac{3}{2}}i+$$
$$+\pi^2 J_0(\pi x)\left[\frac{-\pi i}{(\pi x)^2}+\frac{-\pi^2+(-\pi-\pi^3)i}{(\pi x)^4}+\frac{3\pi^2-2\pi^4+(3\pi-3\pi^3-\pi^5)i}{(\pi x)^6}+\right.$$
$$\left.+\frac{-45\pi^2+12\pi^4-3\pi^6+(-45\pi+27\pi^3-6\pi^5-\pi^7)i}{(\pi x)^8}+\dots\right]+$$
$$+\pi^2 J_1(\pi x)\left[\frac{-1}{\pi x}+\frac{-\pi^2+\pi i}{(\pi x)^3}+\frac{-\pi^2-\pi^4+(-\pi+2\pi^3)i}{(\pi x)^5}+\right.$$
$$\left.+\frac{9\pi^2-3\pi^4-\pi^6+(9\pi-6\pi^3+3\pi^5)i}{(\pi x)^7}+\dots\right].$$

The series are only asymptotic, but by their aid we can tabulate $E_1^0(x)$ and $E_1^1(x)$ beyond the range of convenient application of the ascending series. The function $E_n^s(x)$ is important in the theory of currents in an aerial parallel to the earth. I have not given the details of the reduction referred to above, since that is done in a forthcoming book on aerials and high-frequency networks.

If $I_{mn}(z) \to 0$ as $z \to \infty$, the upper limit in (2) can be taken to be infinity, and then (6) gives a proof of Weber's integral. This is valid for certain non-integral values of m and n, the restriction that m and n should be integers not being essential here.

ON THE MODULUS OF DOUBLY-CONNECTED DOMAINS

By MENAHEM SCHIFFER (*Jerusalem*)

[Received 2 September 1945]

1. Introduction

In this paper conformal properties of doubly-connected domains are studied with the aid of a method of variation. The central role in the theory will be played by the modulus of such a domain; in the present section the definition of the modulus and some of its important properties are recapitulated.

Let D be a domain in the z-plane, bounded by two non-degenerate continua C_1 and C_2 without common points. It is always possible to map D conformally and in a one-to-one manner on a circular annulus $1 < |\zeta| < M$. Hence the circular annulus may be conceived as the canonic type of doubly-connected domains; the problem of conformal representation of given doubly-connected domains on each other reduces, therefore, to the question of conformally mapping circular annuli on each other.

Obviously the linear transformations

$$\zeta = e^{i\psi}z \quad \text{and} \quad \zeta = \frac{e^{i\psi}M}{z} \tag{1}$$

map the annulus $1 \leqslant |z| \leqslant M$ on itself. On the other hand, these are the only possible representations of this annulus upon an annulus in the ζ-plane the interior boundary of which is the circumference $|\zeta| = 1$, i.e. there holds

Lemma 1. *Let $\zeta = f(z)$ be uniform when $1 \leqslant |z| \leqslant M$ and map this annulus univalently upon the annulus $1 \leqslant |\zeta| \leqslant M_1$. Then $M = M_1$, and $f(z)$ is a linear function of the type* (1).

Proof. We may suppose that $f(z)$ maps $|z| = 1$ on $|\zeta| = 1$; for, if not, consider instead of $f(z)$ the function $M_1 f(z)^{-1}$. We form the function

$$h(z) = \log|f(z)| - \frac{\log M_1}{\log M}\log|z|, \tag{2}$$

which is harmonic when $1 < |z| < M$. It vanishes on both boundaries of this annulus and is, therefore, in view of the maximum modulus principle, identically zero. Hence,

$$\log f(z) = \frac{\log M_1}{\log M}\log z + i\psi \quad (\psi \text{ real}). \tag{2'}$$

Since $f(z)$ is uniform and univalent for $1 \leqslant |z| \leqslant M$, we have necessarily

$$\log M = \log M_1 \tag{2''}$$

and, in view of (2′) and (2″),

$$f(z) = e^{i\psi}z, \tag{2'''}$$

which proves the lemma.

Lemma 1 shows that the number M is a characteristic (i.e. uniquely determined) constant of the domain D. Two domains possessing the same constant M may be mapped upon the same annulus and, therefore, upon each other. On the other hand, if two doubly-connected domains can be mapped upon each other, they possess, in view of Lemma 1, the same constant M. This constant M is called the modulus of the doubly-connected domain D. By definition, we have always $M > 1$. With the aid of this notion of the modulus of a domain we may formulate the criterion:

LEMMA 2. *The necessary and sufficient condition that two doubly-connected domains may be mapped upon each other is that their moduli be equal.*

Thus the role of the modulus in the theory of conformal representation is evident. But this same notion is also fundamental in the general theory of functions regular in a circular annulus, i.e. in the theory of Laurent series. This is shown by

THEOREM I. *Let $\zeta = f(z)$ be regular when $1 \leqslant |z| \leqslant M$ and take in this annulus only values which are situated inside a doubly-connected domain* P *of modulus μ, not containing $\zeta = 0$. If z describes the circumference $|z| = 1$, the total increment of the argument of $f(z)$ shall not be zero,*

i.e. $$\Delta \arg f(z) = I\left\{ \oint_{|z|=1} \frac{df(z)}{f(z)} \right\} = 2\pi i n \neq 0.$$

Under these assumptions we have

$$M \leqslant \mu \tag{3}$$

where equality is only possible if $f(z)$ maps the annulus $1 \leqslant |z| \leqslant M$ univalently upon P.

In other words, if $|z| = 1$ is mapped on a curve enclosing the inner boundary of the majorant domain P, then the modulus of P is larger than M.

Proof. Let $\eta = \phi(\zeta)$ be univalent in P and map this domain upon the annulus $1 < |\eta| < \mu$. Hence $\phi[f(z)]$ is regular when $1 < |z| < M$ and, in view of the univalency of $\phi(\zeta)$ and our assumption with respect to $f(z)$, we have

$$\Delta \arg \phi[f(z)] = \pm 2\pi i n$$

if z describes the circumference $|z| = 1$ in the positive sense. The unique determination of $\phi(\zeta)$ still being subject to a linear transformation of the type (1), we may suppose without loss of generality that the argument of $\phi[f(z)]$ increases by a positive multiple of $2\pi i$, i.e. $2\pi i n$ $(n > 0)$. Consider now the function

$$q(z) = z^{-n}\phi[f(z)]. \tag{4}$$

Its argument returns to its initial value if z describes the circumference $|z| = 1$. Further, we obviously have

$$1 \leqslant |q(z)| \leqslant \mu \quad (|z| = 1); \qquad \frac{1}{M^n} \leqslant |q(z)| \leqslant \frac{\mu}{M^n} \quad (|z| = M). \tag{4'}$$

Hence, we have also

$$\Delta \arg [q(z)-r] = \oint\limits_{|z|=1} \frac{dq(z)}{q(z)-r} = 0 \tag{4''}$$

for every value $|r| < 1$, this equality being evident when $r = 0$.

Suppose now that $\mu M^{-n} < 1$ and consider all values r satisfying the inequality

$$\mu M^{-n} < |r| < 1.$$

In view of (4′) and (4″), we have

$$\Delta \arg[q(z)-r] = 0 \tag{4'''}$$

if z describes either the circumference $|z| = 1$ or $|z| = M$. In virtue of the principle of the argument we conclude, therefore, that $q(z)$ does not take these values when $1 \leqslant |z| \leqslant M$. There exists, therefore, a whole annulus $\mu M^{-n} < |q| < 1$ separating the image points of $|z| = 1$ from those of $|z| = M$, which does not contain image points of $1 \leqslant |z| \leqslant M$. This is clearly impossible, since the different image points can be continuously connected. Hence our assumption that $\mu M^{-n} < 1$ leads to a contradiction, whence

$$\mu \geqslant M^n. \tag{5}$$

A more detailed discussion leads without difficulty to the result that

equality in (5) is possible only if $q(z)$ reduces to a constant. In this case we have, therefore,

$$\phi[f(z)] = e^{i\psi}z^n. \tag{5'}$$

Hence $\mu = M$ can only occur when $n = 1$ and

$$f(z) = \phi^{-1}(e^{i\psi}z), \tag{5''}$$

i.e. for a univalent function $f(z)$. This completes the proof of our theorem.

Theorem I may be used to apply the majorant method on Laurent series in the same way as Schwarz's lemma is applied in the case of Taylor series. In this connexion there arises the problem of finding bounds for the modulus of a given doubly-connected domain by means of its elementary geometric properties. For this purpose, I shall establish in the next paragraph a variation formula for the modulus and apply it in the sequel in order to obtain information of the type mentioned.

2. The variation formula for the modulus

Let $\zeta = \phi(z)$ be regular in the doubly-connected domain D and map it in a one-to-one manner upon the annulus $1 < |\zeta| < M$, such that the one boundary continuum C_1 corresponds to the circumference $|\zeta| = 1$, while the other C_2 is mapped on the circumference of radius M. Consider now the function

$$\omega(z) = \frac{\log|\phi(z)|}{\log M}, \tag{6}$$

which is harmonic when $z \subset D$, has on C_1 the boundary value zero, and on C_2 the boundary value 1. We call $\omega(z)$ the *harmonic measure* of C_2 in z with respect to D (**1**). It is well known that $\omega(z)$ admits the integral representation

$$\omega(z) = \frac{1}{2\pi} \oint_{\tilde{C}_2} \frac{\partial g(x;z)}{\partial n_x} ds_x, \tag{7}$$

where $g(x;z)$ is the Green's function of D, and $\tilde{C}_2$ is a closed smooth curve in D which is topologically equivalent to C_2 and separates z from C_2.

Consider now the function

$$z^* = z + \frac{a\rho^2}{z - z_0} \quad (|a| = 1;\ \rho > 0;\ z_0 \subset D), \tag{8}$$

which is obviously univalent in the domain $|z - z_0| > \rho$. If, therefore, z_0 is situated neither on C_1 nor on C_2, ρ may be chosen so small

that both continua lie inside the domain of univalency of (8). Hence they are transformed by this function into other continua C_1^* and C_2^* which determine another doubly-connected domain D^*. Let now $g^*(x;z)$ be the Green's function belonging to D^*; $g^*(x;z)$ may be expressed asymptotically by means of the Green's function $g(x;z)$ of D. For this purpose we define the function $p(x;z)$ which, for fixed $z \subset D$, is analytic in $x \subset D$ and for which

$$g(x;z) = R\{p(x;z)\}. \tag{9}$$

With the help of $p(x;z)$ we may write (3)

$$g^*(x^*;z^*) = g(x;z)+R\{a\rho^2 p'(z_0;x)p'(z_0;z)\}+o(\rho^2). \tag{10}$$

Here x^* and z^* denote the image points of x and z as given by (8); $p'(z_0;y)$ denotes the derivative of $p(z_0;y)$ with respect to its first argument and $o(\rho^2)$ is an expression satisfying the limit condition

$$\lim_{\rho=0}\frac{1}{\rho^2}o(\rho^2) = 0$$

uniformly for every pair $|x-z_0| > \rho_0$, $|z-z_0| > \rho_0$, when x, z belong to an interior partial domain of D.

Combining (10) with (7), we get immediately

$$\omega^*(z^*) = \omega(z)+R\left\{a\rho^2 p'(z_0;z)\cdot\frac{1}{2\pi}\oint_{\tilde{C}_2}\frac{\partial p'(z_0;x)}{\partial n_x}\,ds_x\right\}+o(\rho^2). \tag{7'}$$

By (7) and (9) the function

$$w(z) = \frac{1}{2\pi}\oint_{\tilde{C}_2}\frac{\partial p(z;x)}{\partial n_x}\,ds_x \tag{7''}$$

is obviously analytic when $z \subset D$ and satisfies the equation

$$\omega(z) = R\{w(z)\}. \tag{9'}$$

Hence, remembering (6), one finds

$$w(z) = \frac{\log\phi(z)}{\log M}+iC \quad (C \text{ a real constant}). \tag{11}$$

Thus (7′) may be written in the form

$$\omega^*(z^*) = \omega(z)+R\{a\rho^2 p'(z_0;z)w'(z_0)\}+o(\rho^2) \tag{12}$$

$$= \omega(z)+R\left\{a\rho^2\frac{p'(z_0;z)}{\log M}\,\frac{\phi'(z_0)}{\phi(z_0)}\right\}+o(\rho^2),$$

which gives the variation formula for the harmonic measure.

In view of (6) we have further

$$\oint_{\tilde{C}_2} \frac{\partial \omega(z)}{\partial n_z} ds_z = \frac{1}{\log M} \oint_{\tilde{C}_2} \frac{\partial \log|\phi(z)|}{\partial n_z} ds_z = -\frac{\Delta \arg \phi(z)}{\log M}, \tag{13}$$

i.e., since $\Delta \arg \phi(z) = 2\pi$,

$$(\log M)^{-1} = -\frac{1}{2\pi} \oint_{\tilde{C}_2} \frac{\partial \omega(z)}{\partial n_z} ds_z. \tag{13'}$$

Combining (12) with (13′), we obtain

$$(\log M^*)^{-1} = (\log M)^{-1} - R\left\{a\rho^2 w'(z_0) \frac{1}{2\pi} \oint_{C_2} \frac{\partial p'(z_0; z)}{\partial n_z} ds_z\right\} + o(\rho^2), \tag{14}$$

whence, because of (7″),

$$(\log M^*)^{-1} = (\log M)^{-1} - R\{a\rho^2 w'(z_0)^2\} + o(\rho^2). \tag{14'}$$

Using the formula (11) for $w(z)$, we get finally

$$\log M^* = \log M + R\left\{a\rho^2 \frac{\phi'(z_0)^2}{\phi(z_0)^2}\right\} + o(\rho^2). \tag{15}$$

If the domain D possesses exterior points z_0, we may also consider variations

$$z^* = z + \frac{a\rho^2}{z - z_0} \quad (|a| = 1;\ \rho > 0;\ z_0 \subset \text{exterior of } D). \tag{8'}$$

These are, for small enough ρ, regular and univalent in D and map D upon a domain D^* with modulus M^*. In virtue of lemma 2 of § 1, we have in this case

$$M^* = M. \tag{15'}$$

Formulae (15) and (15′) determine the variation of the modulus with respect to variations (8) and (8′) of the domain D. In the following section a number of applications of these formulae will be made.

3. A distortion theorem for the modulus

Let

$$\zeta = s(z) = z + c_0 + \frac{c_1}{z} + \frac{c_2}{z^2} + \ldots \tag{16}$$

be regular when $|z| > 1$, apart from the pole at infinity, and univalent in the same domain. The function $s(z)$ maps the exterior of the unit circle univalently upon a domain Σ in the ζ-plane. Our object is to study the representation effected by (16) of the circumference $|z| = 1$. For this purpose we determine four different points

$k_\nu = e^{i\kappa_\nu}$ $(\nu = 1, 2, 3, 4)$ on the circumference $|z| = 1$ and consider the arcs $B_1 = \widehat{k_1 k_2}$ and $B_2 = \widehat{k_3 k_4}$, which we suppose to have no common points. By (16) the points of B_i correspond to image points in the ζ-plane, on the boundary of Σ, which form a continuum B_i. The continua B_1 and B_2 may have common points, though the B_i were supposed to be separated.

We consider now the family F of all functions (16) univalent when $|z| > 1$. The fixed arcs B_1 and B_2 on the circumference of the unit circle are mapped by the functions of F on varying image continua B_1 and B_2. It is now easily shown that these image continua cannot be made arbitrarily small. In fact, if we denote the transfinite diameter (outer mapping radius) of B_i by $d(\mathrm{B}_i)$, it may even be shown that the following inequalities hold:

$$d(\mathrm{B}_1) \geqslant \sin^2\tfrac{1}{4}(\kappa_2 - \kappa_1), \qquad d(\mathrm{B}_2) \geqslant \sin^2\tfrac{1}{4}(\kappa_4 - \kappa_3). \tag{17}$$

The proof of these inequalities will be published elsewhere. Here they are only mentioned in order to show that the following extremum problem has a solution:

Among all functions of F, to determine those for which the modulus M of the domain Δ in the ζ-plane, bounded by the continua B_1 *and* B_2, *attains its maximum.*

There exists at least one function solving this problem, since the family F is compact. This function possesses separated image continua B_1 and B_2, since otherwise the modulus of Δ would attain its minimum value 1. On the other hand, it is clear that this maximum M of the modulus must be finite, the modulus of a doubly-connected domain becoming infinite only if one boundary continuum reduces to a point. In our case, this is impossible in view of (17).

After having established the existence of extremal functions and the finiteness of the maximum, we may apply the variation formula of § 2 for the determination of the desired expressions.

Let now $\zeta = \tilde{s}(z)$ be a function (16) of F mapping B_1 and B_2 on continua $\tilde{\mathrm{B}}_1$, $\tilde{\mathrm{B}}_2$, such that the domain $\tilde{\Delta}$ in the ζ-plane, bounded by these continua, possesses the maximum value M of the modulus. Then $\tilde{s}(z)$ maps the complementary arcs $C_1 = \widehat{k_2 k_3}$ and $C_4 = \widehat{k_4 k_1}$ of the circumference $|z| = 1$ on continua $\tilde{\Gamma}_1$ and $\tilde{\Gamma}_2$ of the ζ-plane; these form together with $\tilde{\mathrm{B}}_1$ and $\tilde{\mathrm{B}}_2$ one single continuum, namely the image of the circumference $|z| = 1$ as furnished by $\tilde{s}(z)$. I shall

now characterize $\tilde{s}(z)$ by means of the geometric properties of the continua $\tilde{\mathrm{B}}_i$ and $\tilde{\Gamma}_i$.

I first remark that $\tilde{s}(z)$ maps the domain $|z| > 1$ upon a domain $\tilde{\Sigma}$ which has no exterior points in $\tilde{\Delta}$. For suppose $\zeta_0 \subset \tilde{\Delta}$ to be outside $\tilde{\Sigma}$; then, the function

$$\zeta^*(\zeta) = \zeta + \frac{a\rho^2}{\zeta - \zeta_0} \quad (|a| = 1, \zeta > 0) \tag{8''}$$

will be univalent in $\tilde{\Sigma}$, for small enough ρ, and hence $\zeta^*[\tilde{s}(z)]$ is univalent when $|z| > 1$. The latter function has further the normalization (16) and, accordingly, maps the circular arcs B_1, B_2 on continua B_1^*, B_2^* which, in view of the extremum property of $\tilde{\Delta}$, determine a domain Δ^* with a modulus $M^* \leqslant M$.

On the other hand, we may calculate M^* by means of (15). In fact, we have

$$\log M^* = \log M + R\left\{a\rho^2 \frac{\phi'(\zeta_0)^2}{\phi(\zeta_0)^2}\right\} + o(\rho^2) \tag{18}$$

where $\eta = \phi(\zeta)$ denotes a function mapping $\tilde{\Delta}$ univalently upon the annulus $1 < |\eta| < M$. Since $\phi'(\zeta)$ is not identically zero, we may determine a, ρ, and ζ_0 in such a way that $\log M^* > \log M$, in contradiction to the extremum property of $\tilde{\Delta}$. Hence, the existence of points $\zeta_0 \subset \tilde{\Delta}$ exterior to $\tilde{\Sigma}$ is impossible.

Next, I point out that $\tilde{s}(z)$ is by no means uniquely determined by its extremum property. Every function

$$\psi(\zeta) = \zeta + \gamma_0 + \frac{\gamma_1}{\zeta} + \frac{\gamma_2}{\zeta^2} + \cdots \tag{16'}$$

univalent in $\tilde{\Delta}$ transforms $\tilde{\mathrm{B}}_1$, $\tilde{\mathrm{B}}_2$ into continua $\tilde{\tilde{\mathrm{B}}}_1$, $\tilde{\tilde{\mathrm{B}}}_2$ which bound a domain of equal modulus. Moreover, $\psi[\tilde{s}(z)]$ has also the normalization (16) and, accordingly, furnishes another extremal function. In view of this arbitrariness we may assume from the outset that both continua $\tilde{\mathrm{B}}_1$ and $\tilde{\mathrm{B}}_2$ are circumferences: there remains only the problem of determining the form of the complementary continua $\tilde{\Gamma}_1$ and $\tilde{\Gamma}_2$, taken under this additional condition.

To this end we choose a point $\zeta_0 \subset \tilde{\Delta}$ on $\tilde{\Gamma}_1$. We then determine a partial continuum $\gamma \subset \tilde{\Gamma}_1$, containing ζ_0, of transfinite diameter ρ, and consider the family of all functions

$$\zeta^+(\zeta) = \zeta + \frac{a\rho^2}{\zeta - \zeta_0} + \frac{b\rho^3}{(\zeta - \zeta_0)^2} + \cdots \tag{19}$$

which are univalent in the exterior of γ and regular, with the exception of $\zeta = \infty$. The coefficients $a, b, \ldots$ are bounded by constants which are independent of ζ_0, γ, ρ. The function

$$s^+(z) = \zeta^+[\tilde{s}(z)] = z + l_0 + \frac{l_1}{z} + \ldots \tag{20}$$

is univalent when $|z| > 1$ and has the normalization (16). It maps the arcs B_1 and B_2 on continua B_1^+, B_2^+ which may be obtained from $\tilde{\mathrm{B}}_1$ and $\tilde{\mathrm{B}}_2$ by the transformation (19) of the ζ-plane. They bound a domain Δ^+, the modulus M^+ of which, in view of (15), is given by

$$\log M^+ = \log M + R\left\{a\rho^2 \frac{\phi'(\zeta_0)^2}{\phi(\zeta_0)^2}\right\} + o(\rho^2). \tag{18'}$$

Again, we use the extremum property of M, i.e. $M^+ \leqslant M$, which yields

$$R\left\{a\rho^2 \frac{\phi'(\zeta_0)^2}{\phi(\zeta_0)^2}\right\} + o(\rho^2) \leqslant 0. \tag{21}$$

This inequality holds for any choice of $\zeta_0 \subset \tilde{\Gamma}_1$, of γ and of the function (19).

I now have to make recourse to the following

LEMMA 3 (**2**): *Let Γ be a continuum in the ζ-plane and $\sigma(\zeta) \not\equiv 0$ a function analytic on Γ.*

Further let the inequality

$$R\{a\rho^2\sigma(\zeta_0)\} + o(\rho^2) \leqslant 0 \tag{22}$$

hold for every partial continuum $\gamma \subset \Gamma$ with transfinite diameter ρ and an arbitrary point $\zeta_0 \subset \gamma$ and for each function (19) *univalent in the exterior of γ. Then Γ is an analytic curve with the parametric representation $\zeta(t)$ and satisfying, for properly chosen parameter t, the differential equation*

$$\zeta'(t)^2\sigma[\zeta(t)] = 1. \tag{22'}$$

In virtue of the inequality (21) we may apply this lemma to $\tilde{\Gamma}_1$, whence we infer that $\tilde{\Gamma}_1$ is an analytic curve satisfying the differential equation

$$\zeta'(t)^2 \frac{\phi'[\zeta(t)]^2}{\phi[\zeta(t)]^2} = 1. \tag{23}$$

Obviously $\tilde{\Gamma}_2$ satisfies the same differential equation.

In order to understand the geometrical significance of our result we map the domain $\tilde{\Delta}$ by means of $\eta = \phi(\zeta)$ upon the annulus $1 < |\eta| < M$. The continua $\tilde{\mathrm{B}}_1$ and $\tilde{\mathrm{B}}_2$ correspond to the boundary

circumferences. The curves $\tilde{\Gamma}_i$ correspond to analytic curves $\eta(t) = \phi[\zeta(t)]$ which, in view of (23), satisfy the differential equation

$$\frac{\eta'(t)^2}{\eta(t)^2} = 1 \tag{24}$$

and have, therefore, the parametric representation

$$\eta(t) = Ce^t, \tag{24'}$$

i.e. are radial stretches connecting both circumferences.

I now recall the fact that there are no exterior points of $\tilde{\Sigma}$ in $\tilde{\Delta}$. The image of $\tilde{\Sigma}$ as furnished by $\eta = \phi(\zeta)$ lies, therefore, in the annulus and has there no exterior points. On the other hand, it was shown that the image of $\tilde{\Sigma}$ is bounded by two circular arcs and two radial segments; this, in view of the foregoing fact, is only possible if both radial segments coincide, i.e. if the images of $\tilde{\Gamma}_1$ and $\tilde{\Gamma}_2$ are the two borders of the same radial cut. Thus, we arrive at the following characterization of the extremal domain $\tilde{\Sigma}$:

The extremal domain $\tilde{\Sigma}$ is mapped upon an annulus cut along a radial stretch, if the domain $\tilde{\Delta}$, bounded by the continua $\tilde{\mathrm{B}}_i$, is mapped upon an annulus.

This formulation contains the complete solution of our problem. It permits also an easy calculation of the maximum value M. For this purpose consider the function

$$x(z) = \log \phi[\tilde{s}(z)], \tag{25}$$

which is univalent when $|z| > 1$. Since $\phi[\tilde{s}(z)]$ maps the exterior of the unit circle upon an annulus with a radial cut, $x(z)$ transforms the same domain into the rectangle whose corners are

$$0, \quad 2\pi i, \quad 2\pi i + \log M, \quad \log M.$$

These four corners correspond to the points k_ν $(\nu = 1,..., 4)$, viz. the end-points of the arcs B_1 and B_2.

Now the function mapping the domain $|z| > 1$ upon a rectangle in this way is well known. In fact, let $r(z)$ be univalent when $|z| > 1$ and map this domain upon a rectangle such that k_ν $(\nu = 1,..., 4)$ correspond to 0, $i\omega_1$, $i\omega_1 + \omega_2$, ω_2, with real ω_1 and ω_2. The ratio $\tau = \omega_2/\omega_1$ being fixed uniquely, we have

$$\tau = \frac{1}{2\pi} \log M. \tag{26}$$

Consider now the doubly-connected domain D in the z-plane, bounded

by the circular slits B_1 and B_2. D is mapped by $x = r(z)$ upon the strip $0 < R\{x\} < \omega_2$, as is easily seen by applying Schwarz's symmetry principle. In D, $r(z)$ is univalent but not uniform; to each point $z \subset D$ there corresponds a set of image points $x + 2ni\omega_1$ $(n = 0, \pm 1, \pm 2, \ldots)$. Hence, the function

$$\mu(z) = \exp\left\{\pi\frac{r(z)}{\omega_1}\right\} \tag{27}$$

will be regular and uniform in D. It maps D upon the annulus $1 < |u| < e^{\pi\omega_2/\omega_1}$. The modulus of the domain D bounded by the arcs B_1 and B_2 is therefore equal to $e^{\pi\tau}$. Writing $m(B_1, B_2)$ for this modulus, which, for given B_1 and B_2, is easily calculated, we infer from (26) that

$$\log M = 2\log m(B_1, B_2). \tag{28}$$

THEOREM II. *Let B_1 and B_2 be a pair of arcs on $|z| = 1$ bounding a domain D of modulus $m(B_1, B_2)$. Every conformal representation of $|z| > 1$ transforms them into a pair of continua B_1 and B_2 bounding a domain Δ of modulus $M(\mathrm{B}_1, \mathrm{B}_2)$, such that*

$$\log M(\mathrm{B}_1, \mathrm{B}_2) \leqslant 2\log m(B_1, B_2). \tag{28'}$$

Equality in (28′) holds only if Δ is one of the extremal domains $\tilde{\Delta}$ described above.

Let now C_1 and C_2 be the arcs of the circumference $|z| = 1$ complementary to B_1 and B_2. Obviously the function

$$v(z) = \exp\left\{-\pi i\frac{r(z)}{\omega_2}\right\} \tag{27'}$$

is regular and uniform in the domain D_C of the z-plane, bounded by C_1 and C_2; $v(z)$ maps D_C univalently upon the annulus $1 < |v| < e^{\pi\omega_1/\omega_2}$. Thus we arrive at the following relation between the moduli $m(C_1, C_2)$ and $m(B_1, B_2)$ corresponding to complementary pairs of arcs

$$\log m(B_1, B_2) . \log m(C_1, C_2) = \pi^2. \tag{29}$$

Every conformal representation of $|z| > 1$ transforms B_1, B_2, C_1, C_2 into continua B_1, B_2, Γ_1, Γ_2 respectively, and, in view of (28′) and (29), we obtain

$$\log M(\mathrm{B}_1, \mathrm{B}_2) . \log M(\Gamma_1, \Gamma_2) \leqslant 4\pi^2. \tag{30}$$

If we consider an arbitrary closed curve and divide it into four continua B_1, Γ_1, B_2, Γ_2, we obtain the inequality (30) for the moduli of the domains bounded by B_1, B_2 and Γ_1, Γ_2 respectively. Indeed,

the exterior of the given curve may be mapped upon $|z| > 1$ and the partial continua correspond to four circular arcs on the circumference $|z| = 1$; thus (30) is applicable in this case. In many cases the inequality (30) furnishes a useful means of obtaining bounds for the modulus of a domain bounded by a pair of curves, with the help of another pair of curves bounding a domain with known modulus.

4. The moduli of complementary pairs of curves

Inequality (30) shows an interesting relation between the moduli of two pairs of curves which together form one single continuum. But, while (28′) was the best possible inequality, (30) may still be improved upon, as will be shown in this section.

In virtue of (30) the expression $\log M(\mathrm{B}_1, \mathrm{B}_2).\log M(\Gamma_1, \Gamma_2)$ possesses an upper bound, independent of the special choice of the four partial continua. There exists, therefore, a continuum $\tilde{\mathrm{K}}$, composed of four partial continua $\tilde{\mathrm{B}}_1$, $\tilde{\Gamma}_1$, $\tilde{\mathrm{B}}_2$, $\tilde{\Gamma}_2$, such that the product considered attains its maximum value. I shall determine this maximum with the aid of the variation method.

For this purpose we choose an arbitrary point ζ_0, not belonging to $\tilde{\mathrm{K}}$, and perform the variation (8″) of the ζ-plane. For sufficiently small ρ, the function (8″) is univalent on $\tilde{\mathrm{K}}$ and transforms this curve into a continuum K^*, composed of four partial continua B_1^*, Γ_1^*, B_2^*, Γ_2^*. If $\phi_{\mathrm{B}}(\zeta)$ and $\phi_{\Gamma}(\zeta)$ are the functions mapping the domains bounded by $\tilde{\mathrm{B}}_1$, $\tilde{\mathrm{B}}_2$ and $\tilde{\Gamma}_1$, $\tilde{\Gamma}_2$ on circular annuli, we may express the moduli of the new domains bounded by B_1^*, B_2^* and Γ_1^*, Γ_2^* in terms of these functions by means of (15). In particular we obtain

$$\log M(\mathrm{B}_1^*, \mathrm{B}_2^*).\log M(\Gamma_1^*, \Gamma_2^*) = \log M(\tilde{\mathrm{B}}_1, \tilde{\mathrm{B}}_2).\log M(\tilde{\Gamma}_1, \tilde{\Gamma}_2) + \\ + R\left\{a\rho^2\left[\log M(\tilde{\Gamma}_1, \tilde{\Gamma}_2)\frac{\phi_{\mathrm{B}}'(\zeta_0)^2}{\phi_{\mathrm{B}}(\zeta_0)^2} + \log M(\tilde{\mathrm{B}}_1, \tilde{\mathrm{B}}_2)\frac{\phi_{\Gamma}'(\zeta_0)^2}{\phi_{\Gamma}(\zeta_0)^2}\right]\right\} + o(\rho^2). \tag{31}$$

In view of the extremum property of $\tilde{\mathrm{K}}$, the new product cannot be larger than the product we started out from. On the other hand, we may choose $\operatorname{sgn} a$ arbitrarily. Hence we have necessarily

$$\frac{1}{\log M(\tilde{\mathrm{B}}_1, \tilde{\mathrm{B}}_2)}\frac{\phi_{\mathrm{B}}'(\zeta)^2}{\phi_{\mathrm{B}}(\zeta)^2} + \frac{1}{\log M(\tilde{\Gamma}_1, \tilde{\Gamma}_2)}\frac{\phi_{\Gamma}'(\zeta)^2}{\phi_{\Gamma}(\zeta)^2} = 0 \tag{31′}$$

identically in ζ. I write

$$b = \surd\{\log M(\tilde{\mathrm{B}}_1, \tilde{\mathrm{B}}_2), \qquad c = \surd\{\log M(\tilde{\Gamma}_1, \tilde{\Gamma}_2)\} \tag{31″}$$

and derive from (31′)

$$\frac{1}{b}\log\phi_{\mathrm{B}}(\zeta) = \pm\frac{i}{c}\log\phi_{\Gamma}(\zeta)+\text{constant}. \tag{32}$$

In view of its definition $\log\phi_{\mathrm{B}}(\zeta)$ has a constant real part if ζ moves on $\tilde{\mathrm{B}}_1$ or on $\tilde{\mathrm{B}}_2$. In fact, we have

$$R\{\log\phi_{\mathrm{B}}(\zeta)\} = 0 \quad (\zeta\subset\tilde{\mathrm{B}}_1), \qquad R\{\log\phi_{\mathrm{B}}(\zeta)\} = b^2 \quad (\zeta\subset\tilde{\mathrm{B}}_2). \tag{33}$$

For the same reason

$$R\{\log\phi_{\Gamma}(\zeta)\} = 0 \quad (\zeta\subset\tilde{\Gamma}_1), \qquad R\{\log\phi_{\Gamma}(\zeta)\} = c^2 \quad (\zeta\subset\tilde{\Gamma}_2). \tag{33′}$$

From (32), (33), and (33′), we may deduce

$$I\{\log\phi_{\Gamma}(\zeta)\} = \text{constant} \quad (\zeta\subset\tilde{\mathrm{B}}_i), \qquad I\{\log\phi_{\mathrm{B}}(\zeta)\} = \text{constant} \quad (\zeta\subset\tilde{\Gamma}_i). \tag{33″}$$

If, therefore, we map the domain $\tilde{\Delta}_{\mathrm{B}}$ bounded by $\tilde{\mathrm{B}}_1$, $\tilde{\mathrm{B}}_2$, by means of $\phi_{\mathrm{B}}(\zeta)$, upon an annulus, the images of $\tilde{\Gamma}_1$ and $\tilde{\Gamma}_2$ are radial stretches connecting both circumferences. Since $\phi_{\mathrm{B}}(\zeta)$ is analytic in $\tilde{\Delta}_{\mathrm{B}}$, it follows that $\tilde{\Gamma}_1$ and $\tilde{\Gamma}_2$ are analytic curves. Exactly in the same way it is shown that $\tilde{\mathrm{B}}_1$ and $\tilde{\mathrm{B}}_2$ are likewise analytic curves. Both functions $\log\phi_{\mathrm{B}}(\zeta)$ and $\log\phi_{\Gamma}(\zeta)$ map the continuum $\tilde{\mathrm{K}}$ on a rectangle. $\tilde{\mathrm{K}}$ is composed of analytic curves which do not intersect each other and it divides the ζ-plane into two domains Σ_1 and Σ_2, both being mapped upon rectangles by both functions.

I now give a detailed discussion of these conformal representations. Each continuum $\tilde{\mathrm{B}}_i$ and $\tilde{\Gamma}_i$ contributes one border to the boundary of Σ_1; I shall denote them by $\mathrm{B}_i^{(1)}$ and $\Gamma_i^{(1)}$. The function $\log\phi_{\mathrm{B}}(\zeta)$ maps Σ_1 upon a rectangle with sides of length s_1 and b^2 which correspond to $\mathrm{B}_i^{(1)}$ and $\Gamma_i^{(1)}$ respectively. This follows from (33), since the image of each $\Gamma_i^{(1)}$ is a stretch parallel to the real axis in the strip $0 \leqslant R\{\log\phi_{\mathrm{B}}\} \leqslant b^2$; is_1 is the increment of $\log\phi_{\mathrm{B}}(\zeta)$, if ζ describes the arc $\mathrm{B}_i^{(1)}$.

In the same way $\log\phi_{\Gamma}(\zeta)$ yields the representation of Σ_1 upon a rectangle with sides of length c^2 and s_2, corresponding to the pairs of arcs $\mathrm{B}_i^{(1)}$ and $\Gamma_i^{(1)}$ respectively. According to (32) both rectangles are transformed into each other by a rotation of $\frac{1}{2}\pi$ and a dilatation. We have, therefore,

$$s_1 = \frac{b}{c}c^2 = bc, \qquad s_2 = \frac{c}{b}b^2 = bc. \tag{34}$$

Exactly the same reasoning is now applicable to the complementary domain Σ_2 enclosed by the remaining borders $\mathrm{B}_i^{(2)}$ and $\Gamma_i^{(2)}$.

On the other hand, we note that $\log\phi_{\mathrm{B}}(\zeta)$ increases by $2\pi i$ if ζ describes both borders of $\tilde{\mathrm{B}}_i$, i.e. $\mathrm{B}_i^{(1)}$ and $\mathrm{B}_i^{(2)}$ successively; similarly $\log\phi_{\Gamma}(\zeta)$ increases by the same value if ζ describes $\tilde{\Gamma}_i$. We find, therefore, that Σ_2 is mapped by $\log\phi_{\mathrm{B}}(\zeta)$ upon a rectangle with sides of length $2\pi - s_1$, b^2, and by $\log\phi_{\Gamma}(\zeta)$ upon a rectangle with sides of length c^2, $2\pi - s_2$. Exactly as before, we have

$$2\pi - s_1 = bc = s_1, \qquad 2\pi - s_2 = bc = s_2, \tag{34'}$$

whence

$$s_1 = s_2 = bc = \pi, \tag{34''}$$

and, in view of (31″), eventually

$$\log M(\tilde{\mathrm{B}}_1, \tilde{\mathrm{B}}_2) . \log M(\tilde{\Gamma}_1, \tilde{\Gamma}_2) = \pi^2. \tag{35}$$

Consequently, inequality (30) has been improved to

$$\log M(\mathrm{B}_1, \mathrm{B}_2) . \log M(\Gamma_1, \Gamma_2) \leqslant \pi^2. \tag{35'}$$

It is easily seen that condition (32) and, consequently, equation (35) are satisfied, if $\tilde{\mathrm{B}}_1$, $\tilde{\mathrm{B}}_2$ and $\tilde{\Gamma}_1$, $\tilde{\Gamma}_2$ are complementary pairs of arcs on a circumference. Hence inequality (35′) is the best possible.

5. A related extremum problem

The inequality of the last paragraph is of use in the case of the following problem:

Two curves Γ_1, Γ_2 with the end-points ξ_1, ξ_2 and ξ_3, ξ_4 are given. Consider all domains Δ bounded by two curves B_1 and B_2, connecting ξ_1 with ξ_2 and ξ_3 with ξ_4 without cutting Γ_1 or Γ_2. What is the maximum value of the modulus of Δ?

A complete answer can be given in a somewhat simpler problem of the same type:

Let ξ_ν ($\nu = 1,..., 4$) be four points in the ζ-plane. Determine two continua B_1 and B_2, such that $\xi_1, \xi_2 \subset \mathrm{B}_1$ and $\xi_3, \xi_4 \subset \mathrm{B}_2$ and that the modulus of the domain Δ bounded by these two continua be a maximum.

Again, the existence of extremal continua B_i is assured. Their determination will now be effected by the variation method.

To this end we choose four arbitrary points ζ_ν ($\nu = 1,..., 4$) inside Δ and define the two polynomials

$$z(\zeta) = (\zeta-\zeta_1)(\zeta-\zeta_2)(\zeta-\zeta_3)(\zeta-\zeta_4),$$
$$x(\zeta) = (\zeta-\xi_1)(\zeta-\xi_2)(\zeta-\xi_3)(\zeta-\xi_4). \tag{36}$$

By Lagrange's interpolation formula we have

$$\frac{x(\zeta)}{z(\zeta)} = \sum_{\nu=1}^{4} \frac{x(\zeta_\nu)}{z'(\zeta_\nu)(\zeta-\zeta_\nu)}. \tag{36'}$$

Consider now the variation

$$\zeta^*(\zeta) = \zeta + a\rho^2 \frac{x(\zeta)}{z(\zeta)} \quad (|a| = 1, \rho > 0) \tag{37}$$

of the ζ-plane which, for small enough ρ, is univalent on the continua B_i and which leaves the points ξ_ν unchanged. It transforms the continua B_i into new continua B_i^* bounding a new domain Δ^* the modulus of which is given, in view of (15), (36′), and (37), by

$$\log M^* = \log M + R\left\{a\rho^2 \sum_{\nu=1}^{4} \frac{x(\zeta_\nu)}{z'(\zeta_\nu)} \frac{\phi'(\zeta_\nu)^2}{\phi(\zeta_\nu)^2}\right\} + o(\rho^2). \tag{38}$$

Here again $\phi(\zeta)$ is the univalent function mapping Δ upon an annulus.

In view of the maximum property of M we obviously have $M^* \leqslant M$ and, because of the arbitrariness of $\operatorname{sgn} a$, we obtain

$$\sum_{\nu=1}^{4} \frac{x(\zeta_\nu)}{z'(\zeta_\nu)} \frac{\phi'(\zeta_\nu)^2}{\phi(\zeta_\nu)^2} = 0 \tag{39}$$

for every choice of $\zeta_\nu \subset \Delta$.

If the domain Δ had exterior points, we might have chosen ζ_2, ζ_3, ζ_4 in the exterior of Δ. In virtue of (15′) these points would not have yielded any contribution in the variation formula (38), such that for an arbitrary $\zeta_1 \subset \Delta$,

$$\frac{x(\zeta_1)}{z'(\zeta_1)} \frac{\phi'(\zeta_1)^2}{\phi(\zeta_1)^2} = 0 \tag{39′}$$

would have been satisfied. This, however, is impossible since $\phi'(\zeta) \not\equiv 0$, which proves that Δ has no exterior points.

We return now to (39); this identity holds for every quadruple of points in Δ. It remains valid if all ζ_ν converge towards the same point $\zeta \subset \Delta$. In the limit we obtain, in virtue of a well-known limit formula, the identity

$$\frac{d^3}{d\zeta^3}\left\{x(\zeta) \frac{\phi'(\zeta)^2}{\phi(\zeta)^2}\right\} = 0 \quad (\text{for every } \zeta \subset \Delta). \tag{40}$$

Integrating this identity leads to

$$x(\zeta) \frac{\phi'(\zeta)^2}{\phi(\zeta)^2} = c_1 \zeta^2 + c_2 \zeta + c_3. \tag{40′}$$

If $\zeta = \infty$ is an interior point of Δ, $\phi(\zeta)$ is analytic in $\zeta = \infty$ and admits there an expansion into a series

$$\phi(\zeta) = \alpha_0 + \frac{\alpha_1}{\zeta} + \frac{\alpha_2}{\zeta^2} + \dots.$$

Hence in the vicinity of this point we have

$$x(\zeta)\frac{\phi'(\zeta)^2}{\phi(\zeta)^2} = \frac{\alpha_1^2}{\alpha_0^2} + \frac{\beta_1}{\zeta} + \dots. \tag{41}$$

A comparison of (41) with (40′) yields $c_1 = c_2 = 0$, i.e.

$$x(\zeta)\frac{\phi'(\zeta)^2}{\phi(\zeta)^2} = c_3. \tag{40″}$$

On the other hand, it is always permissible to suppose ∞ to be interior to Δ. For, suppose that, for a given quadruple ξ_ν ($\nu = 1,\dots, 4$), the extremal continua B_i contain the point at infinity. In this case we may perform a linear transformation mapping an interior point of Δ on the point at infinity. The points ξ_ν are transformed into a new quadruple and the B_i into the corresponding extremal continua. For the new quadruple, our assumption leading to (40″) is fulfilled. Since this formula retains its form in a linear transformation, it is proved in full generality.

Consider now the function

$$u(\zeta) = \int\limits_{\xi_1}^{\zeta} \frac{d\zeta}{\sqrt{\{x(\zeta)\}}} \tag{41′}$$

mapping the Riemann surface of $\sqrt{\{x(\zeta)\}}$ upon the u-plane. This function is not uniform in the ζ-plane, but has a uniform inverse function $\zeta(u)$ in the u-plane. Using u as independent variable, we put $\psi(u) = \phi[\zeta(u)]$. In virtue of (40″) and (41′), we have

$$\frac{\psi'(u)^2}{\psi(u)^2} = \frac{\phi'(\zeta)^2}{\phi(\zeta)^2}x(\zeta) = c_3, \tag{42}$$

i.e.

$$\psi(u) = Ce^{\pm u\sqrt{c_3}}. \tag{42′}$$

The continua B_i correspond to continua in the u-plane along which the relation

$$R\{\log\psi(u)\} = R\{\pm u\sqrt{c_3} + \text{constant}\} = \text{constant} \tag{42″}$$

holds. These continua are, therefore, parallel straight lines. The image of B_1 is a straight line containing an infinity of image points of ξ_1 and ξ_2, while the image of B_2 has to pass through an infinity of

image points of ξ_3 and ξ_4. Thus, the full solution of our problem is given by

THEOREM III. *The extremal continua* B_i *are analytic curves in the* ζ*-plane connecting* ξ_1 *with* ξ_2 *and* ξ_3 *with* ξ_4, *which are mapped by the elliptic integral* (41) *on parallel straight lines.*

The numerical value of the maximal modulus can now be easily determined. With

$$\omega_1 = \int\limits_{\xi_1}^{\xi_2} \frac{d\zeta}{\sqrt{\{x(\zeta)\}}}, \qquad \omega_2 = \int\limits_{\xi_3}^{\xi_4} \frac{d\zeta}{\sqrt{\{x(\zeta)\}}}, \tag{43}$$

the function
$$\phi(\zeta) = \exp\left\{-\frac{\pi i}{\omega_1} u(\zeta)\right\} \tag{44}$$

obviously maps the extremal domain Δ upon the annulus

$$1 < |\phi| < e^{I\{\pi\omega_2/\omega_1\}}.$$

Hence, for the value M of the maximal modulus, we arrive at the equation

$$\log M = \pi I\left\{\frac{\omega_2}{\omega_1}\right\}. \tag{45}$$

It is therefore possible to obtain a bound for the modulus of a doubly-connected domain by means of the cross-ratio of two pairs of points, each pair being chosen on a different boundary continuum. Evidently the best upper bound for the modulus will be expressed by means of the modular function.

REFERENCES

1. R. Nevanlinna, *Eindeutige analytische Funktionen* (Berlin, 1936).
2. M. Schiffer, 'A method of variation within the family of simple functions,' *Proc. London Math. Soc.* (2), 44 (1938), 432–49.
3. —— 'Variation of the Green function and the theory of the p-valued functions', *American J. of Math.* 65 (1943), 341–60.

ON THE EXTENDED SPACE OF SEVERAL COMPLEX VARIABLES (I): THE SPACE OF COMPLEX SPHERES

By L. K. HUA (*Peiping*)

[Received 1 October 1945]

In the theory of functions of a complex variable we introduce the point at infinity to make the extended space (Cauchy plane) compact. The procedure of introducing point at infinity depends, in effect, on the group G of linear fractional transformations

$$z^* = (az+b)/(cz+d), \tag{1}$$

where $ad-bc \neq 0$. It is well known that the group G is 'complete' in the sense that any analytic automorph of the extended space is a transformation of the form (1). The aim of the present series of papers is to extend this discussion to the study of functions of several complex variables.

Let G be a group of transformations

$$z_i' = f_i(z_1,..., z_n) \tag{2}$$

of n complex variables. Suppose that G satisfies the assumptions given by Osgood.† The manifold at infinity is introduced by means of the group (2). The totality of finite points and the points at infinity form the extended space $\mathfrak{R} = \mathfrak{R}(G)$. Immediately, we have the problem: 'Is the group G which defines the extended space $\mathfrak{R}(G)$ complete?' More precisely, we seek the group G such that $\mathfrak{R}(G)$ admits no analytic automorph other than those of G.

This important problem has been answered, so far as I am aware, only for two very special cases: (i) the space of the theory of functions, in which the group G is given by

$$z_i^* = (\alpha_k z_k+\beta_k)/(\gamma_k z_k+\delta_k) \quad (\alpha_k \delta_k-\beta_k \gamma_k \neq 0;\ 1 \leqslant i,\, k \leqslant n), \tag{3}$$

and (ii) the extended projective space, in which the group is given by

$$z_i^* = (c_1^{(i)}z_1+...+c_n^{(i)}z_n+c_0^{(i)})/(c_1^{(0)}z_1+...+c_n^{(0)}z_n+c_0^{(0)}), \tag{4}$$

where
$$(c_i^j)_{0 \leqslant i,\, j \leqslant n}$$
is a non-singular matrix.

Recently I established results for several other groups, and they seem to form a complete system in the sense of the structure of

† *Lehrbuch der Funktionentheorie*, II_1 (Teubner, 1929), 293–301.

groups. The present paper contains a proof for the space of complex spheres. I hope to give the other cases in succeeding papers.

In the Lie geometry of hyperspheres in the $(n-1)$-dimensional space, we introduce 'homogeneous coordinates' $(u_1,..., u_n, v_1, v_2)$ to represent a hypersphere with centre $(\xi_1,..., \xi_{n-1})$ and radius R by means of the relations:

$$u_1^2+...+u_n^2-v_1^2-v_2^2 = 0, \tag{5}$$

$$\left.\begin{aligned} u_i &= \rho\xi_i \quad (1 \leqslant i \leqslant n-1),\\ v_2 &= \rho R,\\ u_n &= \tfrac{1}{2}\rho\Big(1-\sum_{i=1}^{n-1}\xi_i^2+R^2\Big),\\ v_1 &= \tfrac{1}{2}\rho\Big(1+\sum_{i=1}^{n-1}\xi_i^2-R^2\Big), \end{aligned}\right\} \tag{6}$$

and, inversely, we have

$$\xi_i = \frac{u_i}{u_n+v_1}, \qquad R = \frac{v_2}{u_n+v_1}.$$

The elements with $u_n+v_1 = 0$ represent improper hyperspheres. The Lie group of the geometry is given by

$$(u_1^*,..., u_n^*, v_1^*, v_2^*) = \rho(u_1,..., u_n, v_1, v_2)F, \tag{7}$$

where F is an $(n+2)$-rowed matrix leaving the quadratic relation (5) invariant. In non-homogeneous coordinates we have

$$\left.\begin{aligned} \xi_i^* &= f_i(\xi_1,..., \xi_{n-1}, R) \quad (1 \leqslant i \leqslant n-1),\\ R^* &= f_n(\xi_1,..., \xi_{n-1}, R). \end{aligned}\right\} \tag{8}$$

If now we extend the geometry to the complex field, we obtain an extended space $\mathfrak{R}$ defined by n complex variables $(\xi_1,..., \xi_{n-1}, R)$, and the new group is obtained from (8) by varying F in the complex field and preserving the relation (5).

For the sake of convenience in the complex field, we can modify our notation slightly. Write

$$\xi_1 = z_1, \quad ..., \quad \xi_{n-1} = z_{n-1}, \quad iR = z_n. \tag{9}$$

Now the homogeneous coordinates of the 'complex sphere' $(z_1,..., z_n)$ are given by

$$x_i = \rho z_i, \quad y_1 = \rho\sum_{i=1}^{n} z_i^2, \quad y_2 = \rho. \tag{10}$$

The transformation takes the form

$$(x_1^*,..., x_n^*, y_1^*, y_2^*) = \rho(x_1,..., x_n, y_1, y_2)F, \tag{11}$$

where F leaves the quadratic relation

$$\sum_{i=1}^{n} x_i^2-y_1y_2 = 0 \tag{12}$$

invariant. Write F as

$$\begin{pmatrix} T & v_1' & v_2' \\ u_1 & a & b \\ u_2 & c & d \end{pmatrix}, \tag{13}$$

where T is an n-rowed matrix and u_1, u_2, v_1, v_2 denote four n-vectors. Now corresponding to (8), we have

$$(z_1^*,\ldots,z_n^*) = \frac{(z_1,\ldots,z_n)T+u_1\sum_{i=1}^n z_i^2+u_2}{(z_1,\ldots,z_n)v_2'+b\sum_{i=1}^n z_i^2+d}. \tag{14}$$

Using (14) instead of (2) I shall prove here that *the group G defined by* (14) *contains all analytic automorphs of the space* $\mathfrak{R}(G)$.

We begin by finding the Laguerre subgroup H of the Lie group G: that is the group of transformations carrying improper points into improper points. More definitely, we look for those (14) which have no variables in the denominator. Consequently $v_2 = 0$, $b = 0$. From

$$\begin{pmatrix} T & v_1' & 0 \\ u_1 & a & 0 \\ u_2 & c & d \end{pmatrix}\begin{pmatrix} I^{(n)} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 \end{pmatrix}\begin{pmatrix} T' & u_1' & u_2' \\ v_1 & a & c \\ 0 & 0 & d \end{pmatrix} = \rho^2\begin{pmatrix} I & 0 & 0 \\ 0 & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 \end{pmatrix} \tag{15}$$

we deduce that

$$TT' = \rho^2 I, \qquad u_1 = 0,$$

$$v_1 = \frac{2}{d}u_2 T', \qquad ad = \rho^2, \qquad c = \frac{u_2 u_2'}{d}.$$

Therefore
$$F = \begin{pmatrix} \rho\Gamma & 2\rho\Gamma u_2'/d & 0 \\ 0 & \rho^2/d & 0 \\ u_2 & u_2 u_2'/d & d \end{pmatrix}, \tag{16}$$

where Γ is orthogonal. F is a product of

$$\begin{pmatrix} \rho & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{17}$$

$$\begin{pmatrix} \Gamma & 2u_2\Gamma & 0 \\ 0 & 1 & 0 \\ u_2 & u_2 u_2' & 1 \end{pmatrix}, \tag{18}$$

and
$$\begin{pmatrix} I & 0 & 0 \\ 0 & 1/d & 0 \\ 0 & 0 & d \end{pmatrix}. \tag{19}$$

Corresponding to (18), we have the mapping

$$\begin{aligned}(z_1^*,\dots,z_n^*) &= (x_1^*,\dots,x_n^*)/y_2^* \\ &= \{(x_1,\dots,x_n)\Gamma+y_2 v_2\}/y_2 = (z_1,\dots,z_n)\Gamma+u_2. \qquad (20)\end{aligned}$$

Corresponding to (17) and (19), we have the mapping

$$(z_1^*,\dots,z_2^*) = \rho(z_1,\dots,z_n). \qquad (21)$$

The group H is generated by (20) and (21).

It follows that any two finite points are equivalent under the transformation (20). For points (i.e. improper complex spheres) at infinity, I make the assertion:

Every point at infinity is carried into a finite point by means of one of the following $n+1$ transformations:

$$z_i^* = -z_i\Big/\Big(\sum_{j=1}^{n} z_j^2\Big) \quad (1 \leqslant i \leqslant n), \qquad (22)$$

and, for a fixed p $(1 \leqslant p \leqslant n)$,

$$\left.\begin{aligned} z_i^* &= z_i\Big/\Big(1-2z_p+\sum_{j=1}^{n} z_j^2\Big) \quad (i \neq p), \\ z_p^* &= \Big(-z_p+\sum_{j=1}^{n} z_j^2\Big)\Big/\Big(1-2z_p+\sum_{j=1}^{n} z_j^2\Big). \end{aligned}\right\} \qquad (23)$$

In fact, write $\qquad \mathfrak{b} = (x_1,\dots,x_n,y_1,y_2).$

Let $\mathfrak{a}$ be a vector $(a_1,\dots,a_n,b_1,b_2)$. Then, evidently

$$\mathfrak{b}^* = -\frac{2\langle\mathfrak{a},\mathfrak{b}\rangle}{\langle\mathfrak{a},\mathfrak{a}\rangle}\mathfrak{a}+\mathfrak{b} \qquad (24)$$

is a transformation of the extended space, where

$$\langle\mathfrak{a},\mathfrak{b}\rangle = \sum_{j=1}^{n} a_j x_j - \tfrac{1}{2}(b_1 y_2+b_2 y_1). \qquad (25)$$

(Actually, this is known as Lie inversion in the geometry of spheres.) Suppose that $\mathfrak{b}$ is a point at infinity: that is, it has $y_2 = 0$. Then we have

$$y_2^* = -\frac{2\langle\mathfrak{a},\mathfrak{b}\rangle}{\langle\mathfrak{a},\mathfrak{a}\rangle} b_2. \qquad (26)$$

Evidently, at least one of the vectors

$$\mathfrak{a} = (0,\dots,0,0,1,1) \qquad (27)$$

and

$$\mathfrak{a} = \big(0,\dots,\underset{p\text{th}}{1},0,\dots,0,1\big) \qquad (28)$$

makes (26) non-vanishing.

Corresponding to (27) we have the transformation

$$z_i^* = \frac{x_i^*}{y_2^*} = \frac{x_i}{-y_1} = -\frac{z_i}{\sum_{j=1}^{n} z_j^2}. \tag{29}$$

Corresponding to (28) we have

$$\mathfrak{b}^* = -(2x_p - y_1)(0,\ldots,1,0,\ldots,0,1)+\mathfrak{b},$$

i.e.

$$x_i^* = x_i \quad (i \neq p), \qquad x_p^* = -(2x_p-y_1)+x_p = -x_p+y_1,$$

$$y_1^* = y_1, \qquad y_2^* = -(2x_p-y_1)+y_2.$$

Then, when $i \neq p$,

$$z_i^* = \frac{x_i^*}{y_2^*} = \frac{x_i}{-2x_p+y_1+y_2}$$

$$= \frac{z_i}{1-2z_p+\sum_{j=1}^{n} z_j^2}, \tag{30_1}$$

and

$$z_p^* = x_p^*/y_2^* = (-x_p+y_1)/(-2x_p+y_1+y_2)$$

$$= \frac{-z_p+\sum_{j=1}^{n} z_j^2}{1-2z_p+\sum_{j=1}^{n} z_j^2}. \tag{30_2}$$

Notice that (30) may be obtained from (29) by means of the transformations

$$z_i = w_i \quad (i \neq p), \qquad z_p = 1-w_p,$$

and $$z_i^* = -w_i^* \quad (i \neq p), \qquad z_p^* = -(1-w_p^*).$$

The Jacobian of the transformation (29) is equal to

$$\Big(\sum_{j=1}^{n} z_j^2\Big)^{-n} \tag{31}$$

and that of (30) is equal to

$$\Big(1-2z_p+\sum_{j=1}^{n} z_j^2\Big)^{-n}. \tag{32}$$

I now assume that $n \geqslant 3$.

Let $$z_i^* = f_i(z_1,\ldots,z_n) \quad (1 \leqslant i \leqslant n) \tag{33}$$

be an analytic mapping carrying the extended space of complex spheres into itself. By a theorem due to Osgood,† the mapping (33) is birational. Consequently, (33) can be written as

$$z_i^* = p_i(z_1,\ldots,z_n)/q(z_1,\ldots,z_n), \tag{34}$$

† Ibid. 299.

where p_i $(1 \leqslant i \leqslant n)$ and q are $n+1$ polynomials without common divisor other than constant.

1. There is a point $(z_1^{(0)},\ldots,z_n^{(0)})$ satisfying

$$p_i(z_1^{(0)},\ldots,z_n^{(0)}) \neq 0$$

and

$$q(z_1^{(0)},\ldots,z_n^{(0)}) \neq 0.$$

The transformation

$$z_i = w_i+z_i^{(0)},$$

of the form (20), converts (34) into a new transformation in which

$$p_i(0,\ldots,0) \neq 0, \qquad q(0,\ldots,0) \neq 0. \tag{35}$$

The transformation

$$z_i^* = f_i\left(\frac{-z_1}{\sum\limits_j z_j^2},\ldots,\frac{-z_n}{\sum\limits_j z_j^2}\right), \tag{36}$$

which is the product of (33) and (22), also maps the extended space on itself. Write

$$z_i^* = \left\{p_i\left(\frac{-z_1}{\sum z_j^2},\ldots,\frac{-z_n}{\sum z_j^2}\right)(\sum z_j^2)^\lambda\right\}\Big/\left\{q\left(\frac{-z_1}{\sum z_j^2},\ldots,\frac{-z_n}{\sum z_j^2}\right)(\sum z_j^2)^\lambda\right\},$$

where λ is the least integer that makes all the numerators and the denominator integral. On account of (35), we find that

$$p_i\left(\frac{-z_1}{\sum z_j^2},\ldots,\frac{-z_n}{\sum z_j^2}\right)(\sum z_j^2)^\lambda$$

and

$$q\left(\frac{-z_1}{\sum z_j^2},\ldots,\frac{-z_n}{\sum z_j^2}\right)(\sum z_j^2)^\lambda$$

are all of degree 2λ.

Consider the Jacobian of (36). Let Δ and Δ_1 be the respective inverses of the Jacobians of (33) and (36). Δ and Δ_1 are polynomials; for, otherwise, there would exist points making the Jacobian vanish. From (36) and (31) we have

$$\Delta_1(z_1,\ldots,z_n) = \Delta\left(\frac{-z_1}{\sum z_j^2},\ldots,\frac{-z_n}{\sum z_j^2}\right)(\sum z_j^2)^n. \tag{37}$$

Since $q(0,\ldots,0) \neq 0$, we have $\Delta(0,\ldots,0) \neq 0$. Consequently, $\Delta_1(z_1,\ldots,z_n)$ is a polynomial of degree $2n$.

Now we may assume, without loss of generality, that p_j and q are polynomials of degree 2λ, that their terms of the highest degree are constant multiples of $(\sum z_j^2)^\lambda$, and that the Jacobian Δ of (34) is a polynomial of degree $2n$ and its terms of highest degree a constant multiple of $(\sum z_j^2)^n$.

2. We decompose the polynomial q into irreducible factors

$$q = q_1^{\lambda_1} \dots q_l^{\lambda_l}. \tag{38}$$

I am going to prove that q_1^n divides the inverse of the Jacobian Δ. Suppose firstly that q_1^2 does not divide $\sum_{j=1}^{n} p_j^2$. The inverse of the Jacobian of the transformations of (29) and (34) is equal to

$$\left\{\sum_{j=1}^{n}\left(\frac{p_j}{q}\right)^2\right\}^n \Delta(z_1,\dots,z_n).$$

It follows that q_1^n divides Δ.

Suppose that q_1^2 divides $\sum_{j=1}^{n} p_j^2$. Without loss of generality, we may assume that q_1 does not divide p_1. The inverse of the Jacobian of the product of (30) and (34) is equal to

$$\left\{1-2\frac{p_1}{q}+\sum_{j=1}^{n}\left(\frac{p_j}{q}\right)^2\right\}^n \Delta(z_1,\dots,z_n).$$

Evidently q_1^n divides $\Delta(z_1,\dots,z_n)$.

Since $\sum z_j^2$ is an irreducible polynomial when $n \geqslant 3$, and Δ is of degree $2n$, we find immediately that $l = 1$ and q_1 is of degree 2. We therefore have

$$q(z_1,\dots,z_n) = \left(a\sum_{j=1}^{n} z_j^2+\dots\right)^\lambda \tag{39}$$

and $$\Delta(z_1,\dots,z_n) = \rho\left(\sum_{j=1}^{n} z_j^2\right)^n+\dots.$$

Consequently

$$\Delta(z_1,\dots,z_n) = \text{constant}\times\{q(z_1,\dots,z_n)\}^{n/\lambda}. \tag{40}$$

By means of a translation (if necessary), we can assume that

$$q(z_1,\dots,z_n) = \left(\sum_{j=1}^{n} z_j^2+c\right)^\lambda. \tag{41}$$

3. The product of (29) and (34) is equal to

$$z_i^* = \frac{p_i}{(\sum p_j^2)/q}. \tag{42}$$

If q does not divide $\sum p_j^2$, there exists a manifold which is mapped

into the point $z_i^* = 0$. This is impossible. Therefore q divides $\sum p_j^2$. By the argument that gave (41), we have

$$\frac{\sum p_j^2}{q} = \Big(a\sum_{j=1}^{n} z_j^2+\sum_{j=1}^{n}\beta_j z_j+\gamma\Big)^{\lambda}. \tag{43}$$

Applying the same argument to the product of (30) and (34), we have immediately

$$q-2p_k+\frac{1}{q}\sum p_j^2 = \Big(\alpha_k\sum_{j=1}^{n} z_j^2+\sum_{j=1}^{n}\beta_{kj} z_j+\gamma_k\Big)^{\lambda}. \tag{44}$$

We suppose that $\lambda > 1$. From (41), (43), (44) we obtain

$$2z_k^* = 1+\left(\frac{\alpha\sum z_j^2+\sum\beta_j z_j+\gamma}{\sum z_j^2+c}\right)^{\lambda}-\left(\frac{\alpha_k\sum z_j^2+\sum\beta_{kj} z_j+\gamma_k}{\sum z_j^2+c}\right)^{\lambda}. \tag{45}$$

Then we have

$$2\frac{\partial z_k^*}{\partial z_l} = \lambda\left(\frac{\alpha\sum z_j^2+\sum\beta_j z_j+\gamma}{\sum z_j^2+c}\right)^{\lambda-1}\frac{\partial}{\partial z_l}\left(\frac{\alpha\sum z_j^2+\sum\beta_j z_k+\gamma}{\sum z_j^2+c}\right)-$$
$$-\lambda\left(\frac{\alpha_k\sum z_j^2+\sum\beta_{kj} z_j+\gamma_k}{\sum z_j^2+c}\right)^{\lambda-1}\frac{\partial}{\partial z_l}\left(\frac{\alpha_k\sum z_j^2+\sum\beta_{kj} z_j+\gamma_k}{\sum z_j^2+c}\right).$$

If there exists a point such that

$$\alpha\sum z_j^2+\sum\beta_j z_j+\gamma = 0, \tag{46}$$

$$\alpha_k\sum z_j^2+\sum\beta_{kj} z_j+\gamma_k = 0 \tag{47}$$

but

$$\sum_{j=1}^{n} z_j^2+c \neq 0,$$

then the point will make the Jacobian vanish. This violates the one-to-one relationship. Thus (46) and (47) imply

$$\sum_{j=1}^{n} z_j^2+c = 0. \tag{48}$$

Consequently, (48) is a linear combination of (46) and (47). Further, $\sum (p_i/q)^2$ cannot be a constant. Thus (47) is a linear combination of (46) and (48). This is impossible when $n \geqslant 3$ because of the independence of $z_1^*,\ldots, z_n^*$.

4. We therefore have $\lambda = 1$. From (41), (43), (44), with some slight modification, we may write (33) as

$$z_k^* = \frac{\sum\beta_{ij} z_j+\gamma_i}{\sum z_j^2+c}. \tag{49}$$

Since q divides $\sum_{j=1}^{n} p_j^2$, we have

$$\sum_{i=1}^{n}\Big(\sum_{j=1}^{n}\beta_{ij} z_j+\gamma_i\Big)^2 = \rho\Big(\sum_{j=1}^{n} z_j^2+c\Big). \tag{50}$$

It follows immediately that

$$\sum_{k=1}^{n} \beta_{ki}\beta_{kj} = p^2\delta_{ij}, \tag{51}$$

$$\sum_{k=1}^{n} \beta_{ki}\gamma_k = 0, \tag{52}$$

and

$$\sum_{k=1}^{n} \gamma_k^2 = c. \tag{53}$$

From (51) we find that

$$\frac{1}{\rho}(\beta_{ki}) = b_{ki}$$

is an orthogonal matrix. From (52) and (53) we get $\gamma_k = 0$, $c = 0$. Thus, by multiplying the transformations of the groups G, (33) now takes the form

$$z_i^* = \frac{\rho \sum_{j=1}^{n} b_{ij} z_j}{\sum_{j=1}^{n} z_j^2}. \tag{54}$$

It belongs evidently to the group G. Therefore, we have established the completeness of the group G when $n \geqslant 3$.

Consequently every element of G is a product of the transformations (20), (21), (22).

I should remark that when $n = 1$ the theorem is well known. When $n = 2$, it is not difficult to establish that the group is not simple, the space being a topological product of two Cauchy planes. More precisely, we use

$$x_1 x_2 - y_1 y_2 = 0$$

instead of (12). The group G may be obtained from

$$z_1^* = \frac{a_1 z_1 + b_1}{c_1 z_1 + d_1}, \qquad z_2^* = \frac{a_2 z_2 + b_2}{c_2 z_2 + d_2},$$

with an additional permutation

$$z_1^* = z_2, \qquad z_2^* = z_1.$$

This is the case known as 'the space of the theory of functions'. Thus we have solved the problem completely.

THE CRITICAL LATTICES OF A CIRCULAR QUADRILATERAL FORMED BY ARCS OF THREE CIRCLES

By KATHLEEN OLLERENSHAW (*Oxford*)

[Received 1 October 1945]

1. Introduction

1.1. VERY recently (**1**) I proved a result concerning lattice points in a non-convex region bounded by arcs of four circles. The proof provided an extremely simple example of a method due to Mordell (**2**): it was entirely self-contained, depending on no theory other than Minkowski's classical theorem on linear forms.

I now notice that similar results (from which the earlier results could in fact have been deduced) hold for another circular quadrilateral bounded by arcs of *three* circles. The proof again is very simple if use is made of some general properties of 'star domains', i.e. regions which are symmetrical about the origin and in which every radius from the origin meets the boundary, a continuous curve, in just one point. The theory of plane star domains has been fully developed by Mahler (**3**). In the application to particular cases his methods may give rise to a large number of difficult extremal problems. It is therefore of some interest that these methods (which are essentially a development of those used by Minkowski for convex regions) should, in the present case, lead to a very simple proof indeed.

The region K_r (Figs. 1–3) considered here is bounded by arcs of three circles $[\pm\frac{1}{2}]$, $[r]$, defined by the equations

$$x^2+y^2 = \pm x, \qquad x^2+y^2 = r^2 \quad (0 < r < 1).$$

More precisely, if P is any point of K_r, then its coordinates (x, y) satisfy at least one of the inequalities

$$x^2+y^2 \leqslant |x|, \tag{1}$$

$$x^2+y^2 \leqslant r^2 \quad (0 < r < 1), \tag{2}$$

i.e. K_r is a closed star domain and consists of the points (x, y) satisfying

$$\min\{(x^2+y^2)^2/x^2,\, (x^2+y^2)/r^2\} \leqslant 1, \tag{3}$$

where the expression in the bracket is homogeneous.

1.2. Consider all those point lattices Λ defined by

$$x = \alpha\xi+\beta\eta, \qquad y = \gamma\xi+\delta\eta, \qquad \Delta \equiv \alpha\delta-\beta\gamma > 0 \tag{4}$$

(where ξ, η run through all integers) which are such that no point of Λ other than the origin is an inner point of these star domains K_r. Such lattices clearly exist: they are called *admissible*. Mahler (**3**) has shown that, for any closed star domain K, there is at least one such admissible lattice Λ of *minimum determinant*, say $\Delta(K)$. Admissible lattices of K of determinant $\Delta(K)$ are called *critical lattices* of K. Mahler has shown further that every critical lattice has at least two independent points on the boundary of K.

As we shall see, the critical lattices of K_r differ according as r lies in one or other of the three ranges

$$0 < r \leqslant \tfrac{1}{2}\surd 2, \qquad \tfrac{1}{2}\surd 2 < r \leqslant \tfrac{1}{2}\surd 3, \qquad \tfrac{1}{2}\surd 3 < r < 1.$$

The results are summarized in three theorems.

THEOREM 1. *When $r \leqslant \frac{1}{2}\surd 2$, the minimum determinant $\Delta(K_r)$ is $r\surd(1-r^2)$ and there are just two critical lattices of K_r, namely L, L', defined respectively by the equations*

$$x = \xi+\eta r^2, \qquad y = \eta r\surd(1-r^2); \tag{5}$$

$$x = \xi-\eta r^2, \qquad y = \eta r\surd(1-r^2). \tag{6}$$

When $r = \frac{1}{2}\surd 2$, then L, L' coincide.

THEOREM 2. *When $\frac{1}{2}\surd 2 \leqslant r \leqslant \frac{1}{2}\surd 3$, the minimum determinant $\Delta(K_r)$ is $2r^3\surd(1-r^2)$ and there are just two critical lattices of K_r, namely the lattice L_1 defined by the equations*

$$\left.\begin{aligned} x &= 4r^2(1-r^2)\xi-r^2\eta \\ y &= r\surd(1-r^2)\{2(2r^2-1)\xi+\eta\} \end{aligned}\right\} \tag{7}$$

and the lattice L_2 which is the reflection of L_1 in the y-axis. When $r = \frac{1}{2}\surd 2$ or $\frac{1}{2}\surd 3$, these two critical lattices coincide.

THEOREM 3. *When $\frac{1}{2}\surd 3 < r < 1$, the minimum determinant $\Delta(K_r)$ is $\frac{1}{2}r^2\surd 3$, and there are an infinity of critical lattices of K_r, namely those given by the equations*

$$\left.\begin{aligned} x &= \xi r\cos\theta+\eta r\cos(\tfrac{1}{3}\pi+\theta) \\ y &= \xi r\sin\theta+\eta r\sin(\tfrac{1}{3}\pi+\theta) \end{aligned}\right\}, \tag{8}$$

where θ is a parameter such that

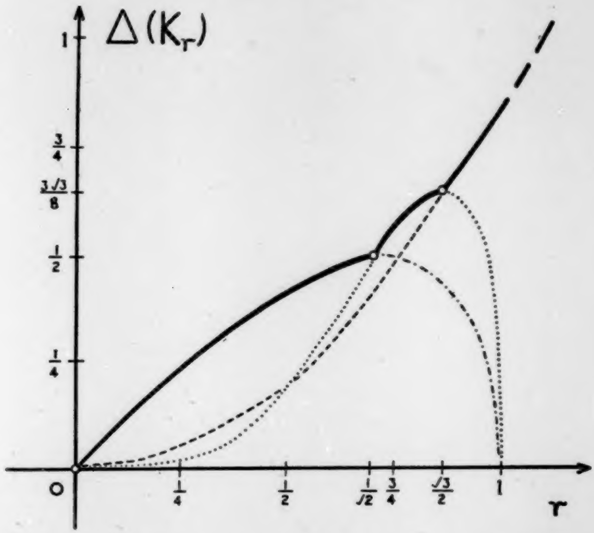
$$\cos^{-1}r \leqslant \theta < \tfrac{1}{3}\pi+\cos^{-1}r.$$

The values of $\Delta(K_r)$ corresponding to all values of r are shown in the accompanying graph. It will be noticed that the curves

$$\Delta(K_r) = r\sqrt{(1-r^2)}, \qquad \Delta(K_r) = 2r^3\sqrt{(1-r^2)}$$

have maxima respectively at the 'hinge' values $r = \frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{3}$. Thus, for all positive values of r less than unity,

$$\Delta(K_r) = \max\{r\sqrt{(1-r^2)},\ 2r^3\sqrt{(1-r^2)},\ \tfrac{1}{2}r^2\sqrt{3}\}. \tag{9}$$



1.3. In § 6 I show that these results can be extended to other regions which contain K_r, and I give examples of such regions when $r \leqslant \frac{1}{2}\sqrt{2}$. This leads to the consideration of *irreducible* regions (discussed and defined in § 7) from which the results of Theorems 1–3 could have been deduced. I construct for every r an irreducible region contained in K_r and having the same 'minimum determinant' as K_r. Such irreducible regions are, presumably, not unique.

1.4. Finally in § 8 I show that Theorems 1–3 are equivalent to the theorem:

THEOREM 4. *If α, β, γ, δ are any real numbers such that*

$$\Delta \equiv \alpha\delta - \beta\gamma > 0,$$

and if f_1, f_2 denote the functions

$$f_1 \equiv (\alpha\xi^2+2b\xi\eta+c\eta^2)^2/(\alpha\xi+\beta\eta)^2, \tag{10}$$

$$f_2 \equiv (a\xi^2+2b\xi\eta+c\eta^2)/r^2, \tag{11}$$

where $\quad a \equiv \alpha^2+\gamma^2, \qquad b \equiv \alpha\beta+\gamma\delta, \qquad c \equiv \beta^2+\delta^2,$

then integers ξ, η not both zero can be found for which

$$\min(f_1,f_2) \leqslant \Delta/\max\{r\sqrt{(1-r^2)},\ 2r^3\sqrt{(1-r^2)},\ \tfrac{1}{2}r^2\sqrt{3}\}. \tag{12}$$

This is a 'best possible' result in the sense that, for given Δ, the expression on the right of (12) is the least number for which this assertion is always true.

2. The region K_r

The circle $[r]$ meets the circles $[\pm\frac{1}{2}]$ in points B, C on the line $y = r\sqrt{(1-r^2)}$. In Figs. 1–3 this line meets the circles $[\pm\frac{1}{2}]$ again in the points* A, D, so that

$$A \text{ is } \{(r^2-1),\ r\sqrt{(1-r^2)}\}, \qquad B \text{ is } \{-r^2,\ r\sqrt{(1-r^2)}\},$$

$$C \text{ is } \{r^2,\ r\sqrt{(1-r^2)}\}, \qquad D \text{ is } \{(1-r^2),\ r\sqrt{(1-r^2)}\}.$$

Further, E is the point $(1,0)$, and A', B',... are the images of A, B,... in O. Thus

$$AC = 1 = BD = OE;$$

hence $OBDE$, $OACE$ are parallelograms, and so

$$OB = ED, \qquad OA = EC.$$

3. Proof of Theorem 1

3.1. Here $0 < r \leqslant \frac{1}{2}\sqrt{2}$ (Fig. 1). The lattice L, defined by (5), is generated by E, C, which correspond to the values $(1,0)$ and $(0,1)$ of (ξ,η) respectively. It has no point other than the origin inside† K_r. For, since $AC = OE$, A is a point of L and the lines $E'A$, EA' produced contain no lattice point inside the circles $[\pm\frac{1}{2}]$. Hence L has no point inside $[\pm\frac{1}{2}]$. Further, $r < 2r\sqrt{(1-r^2)}$, since $r \leqslant 1/\sqrt{2}$, and so no point of the lines $y = \pm 2r\sqrt{(1-r^2)}$ lies in the circle $[r]$; hence L has no point other than O inside $[r]$, and therefore no point other than O inside K_r.

* If $r = \frac{1}{2}\sqrt{2}$, then, of course, A, B coincide, and so also C, D.

† I use 'inside' to mean 'in but not on the boundary of'.

Similarly the lattice L' generated by E', B (or by E, B) has no point other than O inside K_r. By (5), (6), the lattices L, L' both have the determinant $r\sqrt{(1-r^2)}$. It follows that

$$\Delta(K_r) \leqslant r\sqrt{(1-r^2)}. \tag{13}$$

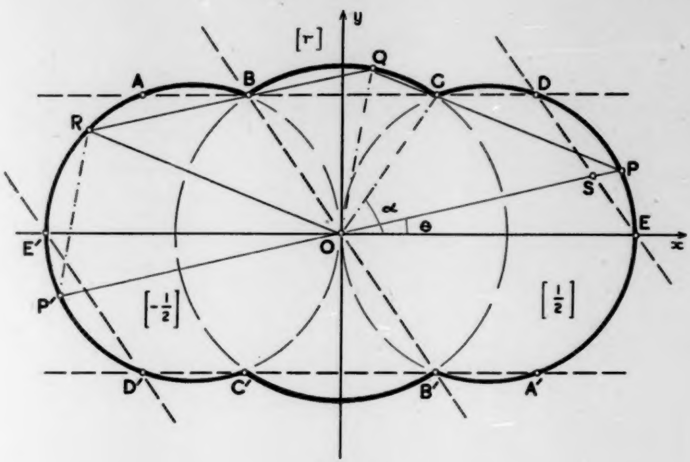


FIG. 1

3.2. It is convenient to prove here an elementary property of star domains.

LEMMA. *If P, Q, R are three points on the boundary of a star domain K such that $OPQR$ is a parallelogram contained in K and of area J, then every lattice Λ of determinant Δ less than J and containing P (or R) has a point not O inside K.*

The proof is immediate. For suppose that Λ is generated by P and a point W on the same side of OP as QR. Then, by hypothesis, the area of the triangle WOP is less than the area of the triangle QOP. Hence the line through W parallel to OP intersects OR, PQ in inner points of these lines and so contains a segment of length OP which is contained in K. Thus either some inner point of this line segment is a lattice point or its point of intersection with OR is a lattice point. Hence, since all inner points of OR are inner points of K (which is not necessarily true of PQ), Λ contains at least one inner point of K other than the origin, which proves the lemma.

3.3. Now, since every critical lattice of K_r contains a point of the boundary of K_r, and since K_r is symmetrical about the coordinate axes, we need consider only those critical lattices of K_r which contain a point of one or other of the circular arcs EC, CB (not B). Consider first the lattices Λ of determinant not greater than $r\sqrt{(1-r^2)}$ which contain E. Since the parallelogram $OEDB$ of area $r\sqrt{(1-r^2)}$ is contained in K_r, every lattice Λ of determinant *less* than $r\sqrt{(1-r^2)}$ and containing E has, by the lemma, a point not O inside K_r. Moreover, every lattice Λ of determinant equal to $r\sqrt{(1-r^2)}$ and containing E contains either both B and D (and then Λ is the lattice L') or some inner point of BD. But all inner points of BD except C are inner points of K_r, and, if Λ contains E, C, then it is the lattice L. Thus either Λ is one of the lattices L, L', or it has a point not O inside K_r.

Similarly, since the parallelogram $ODBE'$ of area $r\sqrt{(1-r^2)}$ is contained in K_r, and since all inner points of BE' are inner points of K_r, every lattice Λ other than L' of determinant not greater than $r\sqrt{(1-r^2)}$ and containing D has a point not O inside K_r. Again, considering the parallelogram $OCAE'$, we see that every lattice Λ other than L of determinant not greater than $r\sqrt{(1-r^2)}$ and containing C has a point not O inside K_r.

3.4. Suppose now that Λ contains an inner point P of the arc ED. Since EC, $E'B$ are the tangents at C, B to the circle $[r]$, the line PC produced meets $[r]$ again in an inner point Q (say) of the arc CB, and the line QB produced meets $[-\frac{1}{2}]$ again in an inner point R (say) of the arc BE'. Now the four points O, P, Q, R are vertices of a parallelogram. For OE is a diameter of $[\frac{1}{2}]$, the quadrilateral $OEPC$ is cyclic, the triangle OCQ is isosceles, and O is the centre of $[r]$, and therefore

$$\angle POE = \tfrac{1}{2}\pi - \angle OEP = \tfrac{1}{2}\pi - \angle OCQ = \tfrac{1}{2}\angle QOC = \angle QBC,$$

i.e. the lines POP', QBR are parallel, and so, similarly, are the lines ROR', QCP.

Now, if OP meets ED in S, then $OP > OS$, and hence, since the triangles OSQ, OSB, OBE each have the same area $\frac{1}{2}r\sqrt{(1-r^2)}$, the area of the parallelogram $OPQR$ is greater than $r\sqrt{(1-r^2)}$. Hence, by the lemma, every lattice Λ of determinant not greater than $r\sqrt{(1-r^2)}$ and containing an inner point of the arc ED has a point not O inside K_r.

3.5. Next, regarding Q as an arbitrary inner point of the arc CB, we see that $OQRP'$ is a parallelogram (with the same area as $OPQR$), since the angles $OP'R$, OBQ, OQB are equal. It follows therefore from the lemma that every lattice Λ of determinant not greater than $r\sqrt{(1-r^2)}$ and containing an inner point of the arc CB has a point not O inside K_r.

3.6. Finally, suppose that Λ contains an inner point P_0 of the arc DC. Then P_0 is above the chord CD and so the triangle P_0OE is greater than the triangle COE. Thus the area of the completed parallelogram OP_0ER_0, where R_0 is a point of the arc $A'B'$, exceeds $r\sqrt{(1-r^2)}$ and is contained in K_r. Hence, by the lemma, every lattice of determinant not greater than $r\sqrt{(1-r^2)}$ and containing P_0 has a point not O inside K_r.

We have thus shown that every lattice of determinant *less* than $r\sqrt{(1-r^2)}$ which contains a point on the boundary of K_r has a point not O inside K_r and hence, by (13),

$$\Delta(K_r) = r\sqrt{(1-r^2)}.$$

Moreover, L, L' are the only lattices of determinant $r\sqrt{(1-r^2)}$ which have no point not O inside K_r but contain points on the boundary of K_r, i.e. L, L' are the only critical lattices of K_r. This completes the proof of Theorem 1.

4. Proof of Theorem 2

4.1. We now consider the critical lattices of K_r when

$$\tfrac{1}{2}\sqrt{2} \leqslant r \leqslant \tfrac{1}{2}\sqrt{3},$$

so that, if $r = \cos\alpha$, i.e. if α denotes the angle EOC (Fig. 2), then

$$\tfrac{1}{6}\pi \leqslant \alpha \leqslant \tfrac{1}{4}\pi.$$

The tangent to $[-\tfrac{1}{2}]$ at B, the point $(-\cos^2\alpha, \cos\alpha\sin\alpha)$, is

$$x\cos 2\alpha - y\sin 2\alpha + \cos^2\alpha = 0.$$

This cuts $[r]$ again in the point Q_1, say $(r\cos\phi_1, r\sin\phi_1)$, where

$$\cos\phi_1\cos 2\alpha - \sin\phi_1\sin 2\alpha + \cos\alpha = 0.$$

Hence
$$\phi_1 = \pi - 3\alpha, \tag{14}$$

and, since $\tfrac{1}{6}\pi \leqslant \alpha \leqslant \tfrac{1}{4}\pi$, we have, by (14), $\tfrac{1}{4}\pi \leqslant \phi_1 \leqslant \tfrac{1}{2}\pi$, and so Q_1 is a point of the arc CY, where Y is $(0, r)$. Since the tangent to $[r]$ at C meets $[\tfrac{1}{2}]$ again in E, the line Q_1C produced meets $[\tfrac{1}{2}]$ again

in a point P_1 of the arc EC. But, if we denote by θ_1 the angle P_1OE, then, as in § 3.4,

$$\theta_1 = \tfrac{1}{2}\pi - \angle OEP_1 = \tfrac{1}{2}\pi - \angle OCQ_1$$
$$= \tfrac{1}{2}\angle Q_1OC = (\tfrac{1}{2}\phi_1 - \tfrac{1}{2}\alpha) = \angle Q_1BC.$$

It follows that OP_1Q_1B is a parallelogram and, by (14), that

$$\theta_1 = \tfrac{1}{2}\pi - 2\alpha.$$

Thus P_1 is the point $\{4r^2(1-r^2), 2r(2r^2-1)\sqrt{(1-r^2)}\}$. We notice too that, when $\alpha = \tfrac{1}{4}\pi$, i.e. when $r = \tfrac{1}{2}\sqrt{2}$, then P_1 is at E and Q_1 is at C.

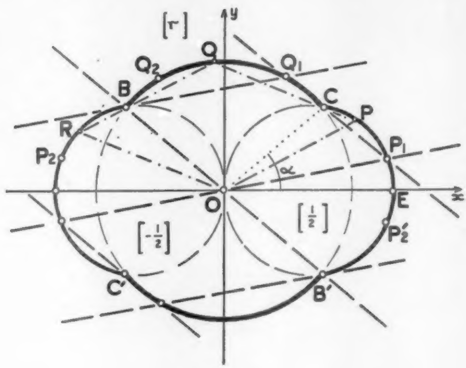


FIG. 2

When $\alpha = \tfrac{1}{6}\pi$, i.e. when $r = \tfrac{1}{2}\sqrt{3}$, then P_1 is at C and Q_1 lies on the y-axis.

The lattice generated by P_1, B contains Q_1 and has the determinant $r^2 \sin 2\alpha = 2r^3\sqrt{(1-r^2)}$: it is the lattice L_1 defined in (7). But L_1 has no point other than the origin inside the region K_r. For no lattice point other than O on the lines BB', P_1Q_1, $P_1'Q_1'$ (produced to infinity in both directions) can lie inside K_r; and, further, since the length of OB is r and the area of OP_1Q_1B is $2r^3\sqrt{(1-r^2)}$, twice the distance between the lines BB', P_1Q_1 is

$$4r^2\sqrt{(1-r^2)} > 4\,.\,\tfrac{1}{2}\sqrt{(1-\tfrac{3}{4})} = 1 = OE.$$

But OE is the greatest distance from O to the boundary of K_r. Thus no lattice point lying beyond P_1Q_1 or $P_1'Q_1'$ lies in K_r and the result follows. Similarly the lattice L_2 which is the reflection of L_1 in the y-axis has no point other than O inside K_r. Hence

$$\Delta(K_r) \leqslant 2r^3\sqrt{(1-r^2)}.$$

4.2. As in the previous case, we need consider only those critical lattices of K_r which contain a point of one or other of the circular arcs EC, CY. Denoting by P_2, Q_2 the reflections of P_1, Q_1 in the y-axis, so that P_2' is the reflection of P_1 in the x-axis, I find it convenient to consider in turn all possible critical lattices contained (i) either of the points P_1, Q_1; (ii) an inner point of either of the arcs P_1C, Q_1Q_2; (iii) an inner point of either of the arcs P_1P_2', CQ_1.

4.3. On considering the parallelograms OP_1Q_1B, OQ_1BP_1' contained in K_r and on noticing that inner points of BQ_1, BP_1' are inner points of K_r, we see, from the lemma, that L_1 is the only admissible lattice of determinant not greater than $2r^3\sqrt{(1-r^2)}$ which contains either of the points P_1, Q_1.

4.4. Suppose next that Λ contains an inner point P of the arc P_1C. Then, since OP_1Q_1B, OCQ_2P_2 are parallelograms and since BQ_1 is the tangent at B to $[-\frac{1}{2}]$, CQ_2 the tangent at C to $[\frac{1}{2}]$, it follows that the line PC produced meets $[r]$ again in an inner point Q of the arc Q_1Q_2, and the line QB produced meets $[-\frac{1}{2}]$ again in an inner point R of the arc BP_2. Moreover, as in §§ 3.4, 4.1, $OPQR$ is a parallelogram.

If now OP meets P_1Q_1 in S, then OP is greater than OS, and hence, since the triangles OSQ, OSB, OBP_1 each have the same area, namely $r^3\sqrt{(1-r^2)}$, the area of the parallelogram $OPQR$ is greater than $2r^3\sqrt{(1-r^2)}$. Thus, by the lemma, every lattice Λ containing an inner point of the arc P_1C and of determinant not greater than $2r^3\sqrt{(1-r^2)}$ has a point not O inside K_r.

Next, regarding Q as an arbitrary inner point of the arc Q_1Q_2, we see that $OQRP'$ is a parallelogram (with the same area as $OPQR$) which is contained in K_r. Hence, by the lemma, every lattice of determinant not greater than $2r^3\sqrt{(1-r^2)}$ and containing an inner point of the arc Q_1Q_2 has a point not O inside K_r.

4.5. Finally, suppose that Λ contains an arbitrary inner point P_0 of the arc P_1P_2' of $[\frac{1}{2}]$. Since $2r \geqslant \sqrt{2} > 1 = OE \geqslant OP_0$, there is a chord of $[r]$ parallel and equal to OP_0 and above OP_0, say R_0Q_0. Since $P_2'C$, P_0Q_0, P_1Q_1 are equal (and parallel) to OQ_2, OR_0, OB respectively, they are each of length r. Hence Q_0 moves from C to Q_1 as P_0 moves from P_2' to P_1. Further, since OR_0 is parallel to P_0Q_0, R_0 moves from Q_2 to B as P_0 moves from P_2' to P_1. Moreover,

if P_0 is on the arc EP'_2, then $OP_0Q_0R_0$ is contained in K_r, since EC is the tangent at C to $[r]$. If P_0 is on the arc EP_1, then, similarly, the parallelogram $OP'_0R_0Q_0$ (of the same area as $OP_0Q_0R_0$) is contained in K_r.

Now, if p is the length of OP_0 so that $0 < p \leqslant 1$, then the area J of $OP_0Q_0R_0$ is given by

$$J = p\sqrt{(r^2-\tfrac{1}{4}p^2)}.$$

Differentiation in p gives

$$dJ/dp = (r^2-\tfrac{1}{2}p^2)(r^2-\tfrac{1}{4}p^2)^{-\frac{1}{2}}.$$

But
$$r^2 \geqslant \tfrac{1}{2} \geqslant \tfrac{1}{2}p^2 > 0,$$

and hence J is an increasing function of p. It follows that the area of $OP_0Q_0R_0$ (and $OP'_0R_0Q_0$) is greater than $2r^3\sqrt{(1-r^2)}$, the area of OP_1Q_1B. Hence, by the lemma and on regarding P_0, Q_0 in turn as arbitrary inner points of the arcs P'_2P_1, CQ_1, we see that every lattice Λ containing an inner point of either of these arcs and of determinant not greater than $2r^3\sqrt{(1-r^2)}$ has a point not O inside K_r.

4.6. We have thus shown that the only possible critical lattice of K_r which contains either C or an inner point of the arcs P'_2P_1C, CQ_1Q_2 is L_1, the lattice generated by P_1, Q_1. By the symmetry of K_r about the coordinate-axes, it follows that L_1, L_2, the lattices defined in Theorem 2, are the only possible critical lattices of K_r when $1/\sqrt{2} \leqslant r \leqslant \tfrac{1}{2}\sqrt{3}$. Hence

$$\Delta(K_r) = 2r^3\sqrt{(1-r^2)},$$

and the proof of Theorem 2 is complete.

5. Proof of Theorem 3

5.1. Lastly $\tfrac{1}{2}\sqrt{3} < r < 1$, so that $0 < \alpha < \tfrac{1}{6}\pi$, where, as usual, $r = \cos\alpha$.

5.2. Mahler (3) has shown that, if a star domain K' contains another star domain K, then $\Delta(K') \geqslant \Delta(K)$. Hence, if at least one critical lattice of K is an admissible lattice of K', then $\Delta(K') = \Delta(K)$, and further every critical lattice of K' is a critical lattice of K. Hence the only critical lattices of K' are those critical lattices of K which have no point other than the origin inside K', since otherwise there would be a critical lattice of K' which was not a critical lattice of K.

5.3. It is well known that the critical lattices of the circle $[r]$ can be defined by the equations

$$\left. \begin{aligned} x &= \xi r \cos\theta + \eta r \cos(\tfrac{1}{3}\pi + \theta) \\ y &= \xi r \sin\theta + \eta r \sin(\tfrac{1}{3}\pi + \theta) \end{aligned} \right\}, \tag{15}$$

where, as usual, ξ, η run through all integers and θ is a parameter. There is no loss of generality in taking

$$\alpha \leqslant \theta < \tfrac{1}{3}\pi + \alpha.$$

5.4. Let $P_1 Q_1$ be the chord of $[r]$ which is equal and parallel to OB and above OB. Then the triangles $B'OP_1$, COQ_1 are equilateral, and so, if θ_1, ϕ_1 are the angles $P_1 Ox$, $Q_1 Ox$, then

$$\theta_1 = \tfrac{1}{3}\pi - \alpha, \qquad \phi_1 = \tfrac{2}{3}\pi - \alpha.$$

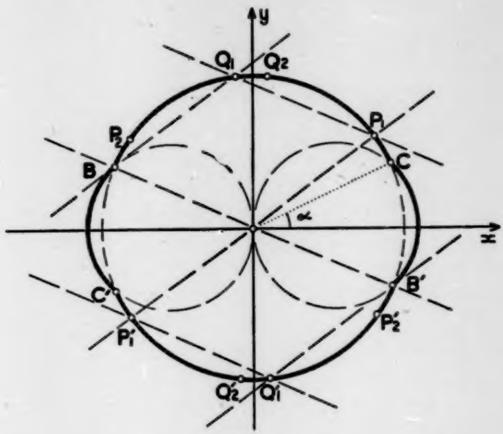


FIG. 3

Hence, since $0 < \alpha < \frac{1}{6}\pi$,

$$\tfrac{1}{6}\pi < \theta_1 < \tfrac{1}{3}\pi, \qquad \tfrac{1}{2}\pi < \phi_1 < \tfrac{2}{3}\pi.$$

Thus P_1 is in the first quadrant and above C, and Q_1 lies in the second quadrant. Moreover, if P_2, Q_2 are the reflections of P_1, Q_1 in the y-axis, then the angles BOP_2, $Q_1 OQ_2$, $P_1 OC$ are all $\frac{1}{3}\pi - 2\alpha$, and so the arcs $C'B$, $P_2 Q_1$, $Q_2 P_1$ of $[r]$, since they subtend angles 2α at O, are equal.

5.5. Now, since $\angle COx = \alpha$ and $\angle Q_2 Ox = \frac{1}{3}\pi + \alpha$, the critical lattices of $[r]$ defined by (15) contain a point of the arc CQ_2 of $[r]$.

Consider first those critical lattices of $[r]$ that contain a point P of the arc CP_1 of $[r]$. By (15) such lattices contain a point Q on $Q_2 Q_1$ and a point R on $P_2 B$, but, since they are critical lattices of $[r]$, they contain no point other than the origin inside $[r]$. Nor do they contain a point inside $[\pm\frac{1}{2}]$. For, firstly, the line PQ produced has no point in common with $[-\frac{1}{2}]$; and, secondly, since PQ is of length r and $\cos\theta \geqslant \cos\theta_1 > \frac{1}{2}$, where θ is the angle POx, the x-coordinate of the point $2P-Q$, i.e. the reflection of Q in the line OP, is

$$r(\cos\alpha+\cos\theta) > \tfrac{1}{2}\sqrt{3}(\tfrac{1}{2}\sqrt{3}+\tfrac{1}{2}) > 1.$$

Hence the line QP produced contains no lattice point which is an inner point of $[\frac{1}{2}]$. Similarly the line QR produced contains no lattice point inside either $[\frac{1}{2}]$ or $[-\frac{1}{2}]$. It follows that the critical lattices of $[r]$ which contain a point of CP_1 (or $Q_1 Q_2$, or $P_2 B$) are admissible lattices of K_r. Hence, by § 5.2, they are *critical* lattices of K_r, since K_r contains $[r]$, and so

$$\Delta(K_r) = \tfrac{1}{2}r^2\sqrt{3}.$$

But the critical lattices of $[r]$ which contain an inner point of the arc $P_1 Q_2$ contain also an inner point of the arc BC' of $[r]$, i.e. an inner point of K_r. They are therefore not admissible lattices of K_r. We have thus shown that the critical lattices of K_r are those defined in (8), and the proof of Theorem 3 is complete.

6. Extensions of Theorem 1

By the argument of § 5.2 it is clear that the region K_r can be enlarged in a variety of ways without altering Theorems 1–3. For example, when $r \leqslant 1/\sqrt{2}$, the results of Theorem 1 still hold if the arc BC of $[r]$ is replaced by an arc (or arcs) Γ through B, C such that any point (x,y) of Γ satisfies the inequalities

$$|x| \leqslant r^2, \qquad r\sqrt{(1-r^2)} \leqslant y \leqslant 2r\sqrt{(1-r^2)}, \qquad x^2+y^2 \geqslant r^2,$$

and the arc $B'C'$ of $[r]$ is replaced by Γ', the reflection of Γ in the x-axis.

Thus Theorem 1 holds for the region K'_r obtained by replacing the circle $[r]$ of K_r by the rectangle

$$|x| \leqslant r^2, \qquad y \leqslant 2r\sqrt{(1-r^2)};$$

or by any of the ellipses

$$x^2r^2(1-r^2)+y^2(a^2-r^4) = a^2r^2(1-r^2),$$

where $$2r^2/\sqrt{3} \leqslant a \leqslant r;$$

or by any pair of ellipses

$$\{b\sqrt{(1-r^2)}-r(1-r^2)\}x^2+r^3(y^2\mp by) = 0,$$

where
$$r/\sqrt{(1-r^2)} \leqslant b \leqslant 2r\sqrt{(1-r^2)}.$$

In the particular case when $b = r/\sqrt{(1-r^2)}$, these ellipses become the circles

$$x^2+y^2 = \pm yr/\sqrt{(1-r^2)},$$

so that any point (x,y) of the new region K'_r satisfies at least one of the inequalities

$$x^2+y^2 \leqslant |x|, \qquad x^2+y^2 \leqslant |yr/\sqrt{(1-r^2)}|.$$

This is in fact the region (referred to in the introduction) for which I first proved results equivalent to those of Theorem 1.

7. Irreducible regions

7.1. This brings us to the consideration of *irreducible** regions, i.e. the smallest regions contained in K_r the critical lattices of which have the same determinant as those of K_r itself. I begin with a formal definition of such regions.

7.2. A star domain K is said to be *irreducible* if, for every other star domain H contained in K, $\Delta(H) < \Delta(K)$. A necessary condition that a star domain K should be irreducible is that every boundary point of K belongs to at least one critical lattice of K. A sufficient condition is that every boundary point belongs to a *free* critical lattice of K, defined as follows:

If Λ is a critical lattice of K of determinant $\Delta(K)$ and with points $\pm P_1$, $\pm P_2$,..., $\pm P_n$ on the boundary of K, and if, corresponding to every $k = 1, 2,..., n$, there is a lattice $\Lambda^{(k)}$ containing points $\pm P_1^{(k)}$, $\pm P_2^{(k)}$,..., $\pm P_n^{(k)}$ arbitrarily near to $\pm P_1$,..., $\pm P_n$ respectively and such that $\pm P_k^{(k)}$ lie inside K but the other points $\pm P_r^{(k)}$ are not inside K, then Λ is said to be a 'free' critical lattice of K.

7.3. Consider first the case when $r \leqslant 1/\sqrt{2}$. Suppose that we replace the arcs CD, $A'B'$, AB, $C'D'$ of the circles $[\pm\frac{1}{2}]$ by the corresponding chords, and call the region so obtained K'_r. Then the determinant of the critical lattices of K'_r is still $r\sqrt{(1-r^2)}$. For the proof is the same as that for the region K_r except that, instead of considering an arbitrary inner point P_0 of the *arc* DC of $[\frac{1}{2}]$ as

* The idea of seeking irreducible regions is due to Mahler. He has recently investigated their properties and I am indebted to him for informing me of the results here stated: these have not yet been published.

in § 3.6, we take P_0 an arbitrary inner point of the *chord* DC. If $P_0 Q_0$ is now drawn parallel and equal to EO, then Q_0 is an inner point of BA and the parallelogram $OEP_0 Q_0$ is contained in K'_r and has the area $r\sqrt{(1-r^2)}$. Moreover, the lattice generated by EP_0 can be seen to have no point other than O inside K'_r, and the result follows. Hence, in addition to the two critical lattices L, L' defined by (5), (6), there are now an infinity of critical lattices generated by E and any arbitrary inner point of the line CD (or AB).

Suppose further that we define the arc of a curve Γ_1 through B and C as the locus of a point U such that, if the line parallel to

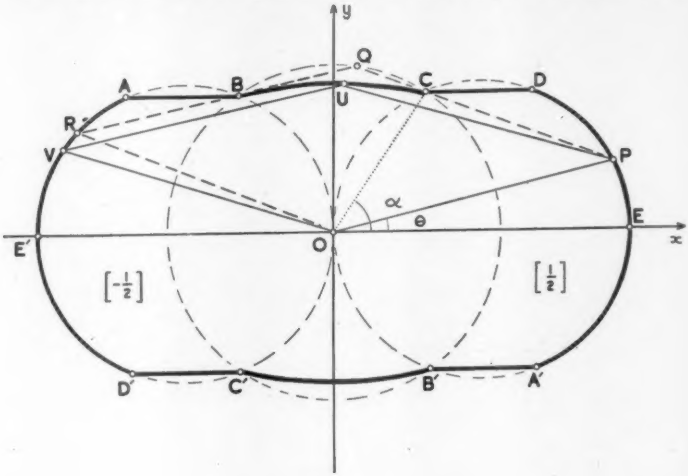


FIG. 4

PO (for variable P on the arc ED) through U cuts the arc $E'A$ of $[-\frac{1}{2}]$ in V, then $OPUV$ is a parallelogram of area $r\sqrt{(1-r^2)}$. Then, since the area of $OPQR$, where Q, R are defined as in § 3.4, is greater than $r\sqrt{(1-r^2)}$, UV lies nearer than RQ to OP. Moreover, since then V must lie outside $OPQR$, it follows that U lies inside $OPQR$, and hence the arc BUC of Γ_1 lies below the circular arc BQC of $[r]$. If now Γ'_1 is the reflection of Γ_1 in the x-axis, and if we denote by G_r the region bounded by the arc $A'ED$ of $[\frac{1}{2}]$, the line DC, the arc CB of Γ_1, the line BA, the arc $AE'D'$ of $[-\frac{1}{2}]$, the line $D'C'$, and the arc $C'B'$ of Γ'_1, then every boundary point of G_r is a base point of at least one critical lattice of G_r.

Moreover, the critical lattices of G_r are *free*. For example, if we consider the critical lattice Λ containing the points P, U, V as shown in Fig. 4, then clearly we can find a lattice Λ' which contains V and points P'', U'' arbitrarily close to P, U, of determinant less than $r\sqrt{(1-r^2)}$, i.e. such that $P''U''$ is nearer than PU to OV, and such that P'' lies inside G_r but U'' does not. Similarly we can find lattices Λ' of determinant less than $r\sqrt{(1-r^2)}$ containing U and points V'', P'' arbitrarily near V, P, and such that V'' lies inside G_r but P'' does not. So for the other points of Λ on the boundary of G_r. In the same way every critical lattice of G_r is free, and hence G_r is an irreducible region.

The required curve Γ_1 is the bicircular quartic

$$(x^2+y^2)^2 = x^2+4yr\sqrt{(1-r^2)}-4r^2(1-r^2). \tag{16}$$

This is obtained by eliminating x', y', t between the four equations

$$y' = x't+(1+t^2)r\sqrt{(1-r^2)}, \qquad x'^2+y'^2+x' = 0,$$
$$x = x'+1/(1+t^2), \qquad y = y'+t/(1+t^2),$$

where U as defined above is the point (x,y), V is (x',y'), P is $(\cos^2\theta, \cos\theta\sin\theta)$, t is $\tan\theta$.

7.4. I now construct an irreducible region when $1/\sqrt{2} \leqslant r \leqslant \frac{1}{2}\sqrt{3}$ using the notation of § 4.

First we notice that, if P is an arbitrary point of the arc CQ_1 of $[r]$ and if $OPQR$ is the parallelogram obtained by rotating OCQ_2P_2 about the vertex O, then, as P moves from C to Q_1 along $[r]$, Q traces the arc Q_2B of $[r]$ and R traces the arc P_2P_1' of the circle with centre O and radius OP_2, say $[OP_2]$. Moreover, QR cuts $[r]$ in Q and some point between B and $(-r,0)$. Hence inner points of QR are inner points of the new region obtained from K_r by replacing the arcs P_2P_1' of $[-\frac{1}{2}]$, $P_2'P_1$ of $[\frac{1}{2}]$ by arcs of $[OP_2]$. Thus, by the lemma, the determinant of the critical lattices of the new region so obtained is still $2r^3\sqrt{(1-r^2)}$ and every point of the arcs CQ_1, Q_2B, $C'Q_1'$, $Q_2'B'$ of $[r]$, and $P_2'P_1$, P_2P_1' of $[OP_2]$ form a base point of a critical lattice of this region.

If, further, we replace the arc Q_1Q_2 of $[r]$ by the arc of the bicircular quartic Γ_2,

$$(x^2+y^2)^2 = x^2+8yr^3\sqrt{(1-r^2)}-16r^6(1-r^2), \tag{17}$$

obtained from Γ_1 of the previous paragraph by writing $2r^3\sqrt{(1-r^2)}$ for $r\sqrt{(1-r^2)}$ in (16), and replace the arc $Q_1'Q_2'$ of $[r]$ by a similar

arc of Γ_2', the reflection of Γ_2 in the x-axis, then, as before, every point of the arcs $P_1 C$, BP_2, $P_1' C'$, $B' P_2'$ of $[\pm\frac{1}{2}]$, and of the arcs $Q_1 Q_2$ of Γ_2, $Q_1' Q_2'$ of Γ_2' is a base point of a critical lattice, of determinant $2r^3\sqrt{(1-r^2)}$, of the new region F_r shown in Fig. 5 and formed

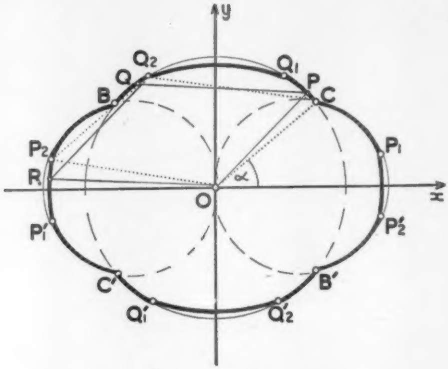


FIG. 5

by arcs of the circles $[\pm\frac{1}{2}]$, $[OP_2]$, $[r]$, and arcs of the curves Γ_2, Γ_2'. Moreover, as before, the critical lattices of F_r are free. Hence the region F_r is irreducible.

7.5. We notice that, when $r = \frac{1}{2}\sqrt{2}$, then P_2, P_1' coincide at $(-1, 0)$; Q_2, Q_1 are at B, C; and the regions F_r, G_r are identical.

When $r = \frac{1}{2}\sqrt{3}$, then Q_1, Q_2 coincide at $(0, r)$, P_1 is at C, P_2' is at B', and F_r is the circle $[r]$. This is to be expected, since the simplest irreducible region of K_r when $r \geqslant \frac{1}{2}\sqrt{3}$ is the circle $[r]$ itself.

8. Proof of Theorem 4

Consider now the functions f_1, f_2 defined in (10), (11). They can be written as

$$(x^2+y^2)^2/x^2, \qquad (x^2+y^2)/r^2,$$

where x, y are points of the lattice Λ defined by (4). Thus $\min(f_1, f_2)$ is a homogeneous function of the second degree in x, y and there is a constant k, independent of α, β, γ, δ, for which integers ξ, η, not both zero, can be found to satisfy

$$\min(f_1, f_2) \leqslant k\Delta. \tag{18}$$

When $k\Delta = 1$, then (18) defines the region K_r and $\Delta(K_r)$ is the greatest value of Δ for which every lattice of determinant Δ has

a point not O in K_r. Hence $k = 1/\Delta(K_r)$ is the best possible, i.e. the least, value of k in (18). This establishes Theorem 4, since, by (9),

$$\Delta(K_r) = \max\{r\sqrt{(1-r^2)},\, 2r^3\sqrt{(1-r^2)},\, \tfrac{1}{2}r^2\sqrt{3}\}.$$

The various critical lattices lead to 'critical pairs of forms' of f_1, f_2, i.e. forms for which the equality sign is necessary in (12). For example, when $r \leqslant 1/\sqrt{2}$, if in (4) we choose one of the two particular matrices

$$\begin{Bmatrix} \alpha & \beta \\ \gamma & \delta \end{Bmatrix} \equiv \begin{Bmatrix} 1 & \pm r^2 \\ 0 & r\sqrt{(1-r^2)} \end{Bmatrix},$$

so that Λ is one of the lattices L, L', then f_1, f_2 are

$$f_1' \equiv (\xi^2 \pm 2\xi\eta r + \eta^2 r^2)^2/(\xi \pm \eta r^2)^2,$$

$$f_2' \equiv (\xi^2 \pm 2\xi\eta r + \eta^2 r^2)/r^2,$$

where either the positive or the negative sign is to be taken throughout; and these are the critical pairs of forms.

My thanks are due to Professor L. J. Mordell, Dr. K. Mahler, and Mr. T. W. Chaundy for their continued help.

REFERENCES

1. K. Ollerenshaw. See above, pp. 93–8.
2. L. J. Mordell, *Proc. London Math. Soc.* (2), 48 (1945), 339–90.
3. K. Mahler, ibid. (2), 49 (1946), 128–83.

ON INFINITE SETS OF HOMOGENEOUS LINEAR EQUATIONS IN HILBERT SPACE

By S. H. HILDING (*Stockholm*)

[Received 24 March 1946]

THE purpose of this note is to give a criterion when the infinite set of equations

$$\sum_{q=1}^{\infty} \frac{x_q}{\lambda_n - \mu_q} = 0 \quad \begin{array}{l} (1 \leqslant n < \infty) \\ (\lambda_n \neq \mu_q \text{ for all } n,\, q) \end{array}$$

has no solution of the class L_2 $\left(\text{i.e. such that } \sum_1^{\infty} |x_q|^2 < \infty\right)$ other than the trivial $x_q = 0$ $(1 \leqslant q < \infty)$.

The set $\{f_n\}$ in Hilbert space H is *complete* if

$$(x, f_n) = 0 \quad (1 \leqslant n < \infty) \tag{1}$$

has only the solution $x = 0$. The set $\{f_n\}$ is *closed* if to every element g of H and every given $\epsilon > 0$ we can find a sequence of coefficients $\{c_n\}$ and a number M such that

$$\left\| g - \sum_1^M c_n f_n \right\| < \epsilon \tag{2}$$

is satisfied.

According to a well-known theorem $\{f_n\}$ is closed if and only if it is complete. With a complete orthonormal set $\{h_n\}$ as coordinate-system in H, (1) can be written

$$\sum_{q=1}^{\infty} x_q a_{nq} = 0 \quad (1 \leqslant n < \infty), \tag{3}$$

where $x_q = (h_q, x)$ and $a_{nq} = (f_n, h_q)$.

The set of equations (3) has thus no solution of the class L_2 other than $x_q = 0$ $(1 \leqslant q < \infty)$ if (2) is satisfied. But (2) is satisfied if and only if there are numbers $\{N_p\}$ and a sequence $\{c_{np}\}$ such that

$$\left\| h_p - \sum_{n=1}^{N_p} c_{pn} f_n \right\| < \epsilon_p \quad (1 \leqslant p < \infty) \tag{4}$$

for arbitrary positive $\{\epsilon_p\}$. The necessity of (4) for (2) is evident, and the sufficiency follows from the inequality $\|f+g\| \leqslant \|f\| + \|g\|$; for, putting $\gamma_p = (g, h_p)$ and taking P so large that

$$\sum_{P+1}^{\infty} |\gamma_p|^2 < \tfrac{1}{4}\epsilon^2$$

and $\{\epsilon_p\}$ such that $$\|g\|^2 \sum_1^\infty |\epsilon_p|^2 < \tfrac{1}{4}\epsilon^2,$$
we have

$$\Big\|g-\sum_{n=1}^{M} c_n f_n\Big\| \leqslant \Big\|\sum_1^P \gamma_p h_p - \sum_1^M c_n f_n\Big\| + \Big\|\sum_{P+1}^{\infty} \gamma_p h_p\Big\|$$

$$< \sum_1^P |\gamma_p| \cdot \Big\|h_p - \sum_{n=1}^{N_p} c_{pn} f_n\Big\| + \tfrac{1}{2}\epsilon < \Big(\sum_1^P |\gamma_p|^2 \sum_1^P |\epsilon_p|^2\Big)^{\frac{1}{2}} + \tfrac{1}{2}\epsilon < \epsilon,$$

where $$c_n = \sum_{p=1}^{P} \gamma_p c_{pn} \quad \text{and} \quad M = \max_{1 \leqslant p \leqslant P} N_p.$$

Now (4) can be written

$$\epsilon_p^2 > \Big\|h_p - \sum_{n=1}^{N_p} c_{np} \sum_{q=1}^{\infty} a_{nq} h_q\Big\|^2$$

$$= \sum_{q=1}^{N_p} \Big|\sum_{n=1}^{N_p} c_{np} a_{nq} - \delta_{pq}\Big|^2 + \sum_{q=N_p+1}^{\infty} \Big|\sum_{n=1}^{N_p} c_{np} a_{nq}\Big|^2,$$

where as usual $$\delta_{pq} = \begin{cases} 1 & \text{if } p = q, \\ 0 & \text{if } p \neq q. \end{cases}$$

If the determinant $\det\limits_{N_p} |a_{nq}|$ of order N_p is not zero, the set of equations

$$\sum_{n=1}^{N_p} c_{np} a_{nq} = \delta_{pq} \quad (1 \leqslant q \leqslant N_p)$$

has a single solution, say $c_{np} = C_{np}$ $(1 \leqslant n \leqslant N_p)$. Then (4) is satisfied when, for every p,

$$\lim_{N_p \to \infty} \sum_{q=N_p+1}^{\infty} \Big|\sum_{n=1}^{N_p} C_{np} a_{nq}\Big|^2 = 0,$$

and in that case (3) has only the solution $x_q = 0$ $(1 \leqslant q < \infty)$ of the class L_2. In particular, if $a_{nq} = \dfrac{1}{\lambda_n - \mu_q}$, we get

$$D_{N_p} = \det_{N_p} \Big|\frac{1}{\lambda_n - \mu_q}\Big| = \frac{\prod (\lambda_r - \lambda_s) \prod (\mu_r - \mu_s)}{\prod \prod (\lambda_m - \mu_n)},$$

where

$$1 \leqslant s < r \leqslant N_p, \qquad 1 \leqslant m \leqslant N_p, \qquad 1 \leqslant n \leqslant N_p.$$

We can disregard the cases $\lambda_r = \lambda_s$ when $r \neq s$ and $\mu_p = \mu_q$ when $p \neq q$, which are trivial. Moreover, when $N_p > p$, $C_{np} = A_{np}/D_{N_p}$ $\Big($if, as usual, $\det\limits_{N} |a_{nq}| = \sum\limits_{i=1}^{N} a_{ik} A_{ik}\Big)$, and consequently

$$\sum_{n=1}^{N_p} C_{np}\, a_{nq} = \frac{1}{D_{N_p}} \sum_{n=1}^{N_p} A_{np}\, a_{nq}$$

$$= \frac{1}{D_{N_p}} \begin{vmatrix} (\lambda_1-\mu_1)^{-1} & . & (\lambda_{N_p}-\mu_1)^{-1} \\ (\lambda_1-\mu_2)^{-1} & . & . \\ . & . & . \\ (\lambda_1-\mu_{p-1})^{-1} & . & . \\ (\lambda_1-\mu_q)^{-1} & . & . \\ (\lambda_1-\mu_{p+1})^{-1} & . & . \\ . & . & . \\ (\lambda_1-\mu_{N_p})^{-1} & . & (\lambda_{N_p}-\mu_{N_p})^{-1} \end{vmatrix} = \pm \prod_{\substack{i=1 \\ i\neq p}}^{N_p} \frac{(\mu_i-\mu_q)}{(\mu_i-\mu_p)} \prod_{i=1}^{N_p} \frac{(\lambda_i-\mu_p)}{(\lambda_i-\mu_q)}.$$

Thus we have the following theorem:

THEOREM. *The infinite set of equations*

$$\sum_{q=1}^{\infty} \frac{x_q}{\lambda_p-\mu_q} = 0 \quad (p = 1,2,3,\ldots) \quad \begin{matrix} (\lambda_p \neq \mu_q \text{ for all } p,\, q) \\ (\mu_p \neq \mu_q \text{ for } p \neq q) \end{matrix}$$

has no solution such that $\sum_1^{\infty} |x_q|^2 < \infty$ *except* $x_q \equiv 0$, *when*

$$\lim_{N\to\infty} \sum_{q=N+1}^{\infty} |R_q(N)|^2 = 0 \quad \textit{for} \quad p = 1,2,3,\ldots,$$

where

$$R_q(N) = \prod_{\substack{i=1 \\ i\neq p}}^{N} \frac{(\mu_i-\mu_q)}{(\mu_i-\mu_p)} \prod_{i=1}^{N} \frac{(\lambda_i-\mu_p)}{(\lambda_i-\mu_q)}.$$

In several cases, this criterion is easily used.

Examples. $\lambda_p = p-\lambda, \qquad \mu_q = q$

gives $$\sum_{q=1}^{\infty} \frac{x_q}{p-\lambda-q} = 0 \quad (1 \leqslant p < \infty),\ \lambda \text{ not an integer.} \tag{A}$$

Here

$$R_q(N) = \left| \frac{\Gamma(q)\Gamma(q+\lambda-N)\Gamma(N-p-\lambda+1)}{(q-p)\Gamma(q-N)\Gamma(p)\Gamma(N+1-p)\Gamma(q+\lambda)\Gamma(-p-\lambda)} \right|$$

$$= O\left[\frac{(q-N)^{\lambda}(N-p+1)^{-\lambda}}{q^{\lambda}(q-p)}\right]$$

and the criterion is

$$\lim_{N\to\infty} \frac{1}{N^{2\lambda}} \sum_{q=N+1}^{\infty} \left(1-\frac{N}{q}\right)^{2\lambda} \frac{1}{q^2} = 0.$$

Thus (A) *has only the* L_2*-solution* $x_q \equiv 0$ *when* $\lambda > -\frac{1}{2}$ [cf. (**1**)].

With $\lambda_p = p$, $\mu_q = aq$ (a irrational) we get

$$\sum_{q=1}^{\infty} \frac{x_q}{p-aq} = 0 \quad (1 \leqslant p < \infty). \tag{B}$$

Here

$$R_q(N) = \left|\frac{\Gamma(q)\Gamma(N+1-ap)\Gamma(aq-N)}{(q-p)\Gamma(q-N)\Gamma(p)\Gamma(N+1-p)\Gamma(-ap)\Gamma(aq)}\right|$$
$$= O\left[\frac{N^{p(1-a)}\Gamma(q)\Gamma(aq-N)}{q\Gamma(aq)\Gamma(q-N)}\right].$$

If $a > 1$, $\qquad \Gamma(q)\Gamma(aq-N) < \Gamma(aq)\Gamma(q-N)$,

and thus $\qquad \lim_{N\to\infty} \sum_{q=N+1}^{\infty} |R_q(N)|^2 = 0.$

For $a > 1$ the only L_2-solution of (B) *is thus $x_q \equiv 0$* [cf. (2)].

As a third example, let us consider

$$\sum_{q=1}^{\infty} \frac{x_q}{p-d_p-q} = 0 \quad (1 \leqslant p < \infty), \qquad d_n = O(1). \tag{C}$$

Here $\qquad R_q(N) = \left|\frac{d_p P_N(q)}{q-p+d_p}\right|,$

where

$$P_N(q) = \prod_{\substack{i=1\\ i\neq p}}^{N} \frac{1+(d_i/p-i)}{1+(d_i/q-i)} = \exp\left[\sum_{\substack{i=1\\ i\neq p}}^{N} \left\{\log\left(1+\frac{d_i}{p-i}\right)-\log\left(1+\frac{d_i}{q-i}\right)\right\}\right]$$
$$= O\left[\exp\left\{\sum_{\substack{i=1\\ i\neq p}}^{N} d_i\left(\frac{1}{p-i}-\frac{1}{q-i}\right)\right\}\right]$$
$$= O\left[\exp\left\{-\sum_{i=p+1}^{N} d_i\left(\frac{1}{i-p}+\frac{1}{q-i}\right)\right\}\right],$$

when $d_i \geqslant 0$.

Thus $P_N(q) = O(1)$. If $-\theta \leqslant d_i < 0$, we get

$$P_N(q) = O\left[\exp\left\{\theta\log(N-p)+\theta\log\frac{q-N}{q-p}\right\}\right] = O\left[N^\theta\left(1-\frac{N}{q}\right)^\theta\right].$$

Thus

$$\sum_{q=N+1}^{\infty} |R_q(N)|^2 = O\left(\sum_{N+1}^{\infty}\frac{1}{q^2}\right) = o(1) \quad \text{when} \quad d_i \geqslant 0,$$
$$= O\left(N^{2\theta}\sum_{N+1}^{\infty}\frac{1}{q^2}\right) = O(N^{2\theta-1}) = o(1)$$
$$\text{when} \quad 0 > d_i \geqslant -\theta > -\tfrac{1}{2}.$$

It may be interesting to compare (C) with

$$\sum_{q=-\infty}^{\infty} \frac{x_q}{p-d_p-q} = 0 \quad (-\infty < p < \infty), \qquad d_n = O(1). \tag{C$'$}$$

For, *when* $-\frac{1}{2} < \theta \leqslant d_n \leqslant K$ (K *arbitrary*), *the only* L_2*-solution of* (C) *is* $x_q \equiv 0$, *and, when* $-\frac{1}{2} \leqslant d_n \leqslant 0$, *the only* L_2*-solution of* (C$'$) *is* $x_q \equiv 0$. The latter statement follows from the fact that

$$\{\exp i(n-d_n)x\}$$

is a complete system in $L_2(-\pi,\pi)$, when $|d_n| \leqslant \frac{1}{4}$ [cf. (**3**)], and so in every interval of length $\frac{1}{2}$; for $a_{nq} = (f_n, h_q)$ in (3) corresponds in $L_2(-\pi,\pi)$ to

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(n-d_n-q)x}\,dx = \frac{(-1)^q \sin\pi(n-d_n)}{\pi(n-d_n-q)} \quad (-\infty < n,\, q < \infty).$$

The length of the interval $\frac{1}{2}$ is 'best possible' for (C$'$) [cf. (**3**)], and the constant $\frac{1}{2}$ is also 'best possible' for (C), because (A), which is evidently a special case of (C), has, for $-1 < \lambda < 0$, the solution [cf. (**1**)]

$$x_q = \frac{\Gamma(q+\lambda)}{\Gamma(q)} \quad (1 \leqslant q < \infty)$$

and this is an L_2-solution for every $\epsilon > 0$, when $\lambda < -(\frac{1}{2}+\epsilon)$.

The set of equations

$$\sum_{q=1}^{\infty} \frac{x_q}{p-\mu_q} = 0 \quad (1 \leqslant p < \infty) \tag{D}$$

may also be treated in a similar manner to (C) under different hypotheses on $\{\mu_q\}$ [cf. (**4**)].

REFERENCES

1. Linfoot and Shepherd, 'On a Set of Linear Equations': *Quart. J. of Math.* (Oxford), 10 (1939), 1–10.
2. Persson, *Sur une classe de systèmes d'équations linéaires à une infinité d'inconnues* (Uppsala, 1938, thèse), 22.
3. Levinson, 'On the Closure of $\{\exp i\lambda_n x\}$': *Duke Math. J.* 2 (1936), 511–16.
4. Persson, loc. cit. 49.

THE NON-LINEAR DIFFERENCE-DIFFERENTIAL EQUATION

By E. M. WRIGHT (*Aberdeen*)

[Received 25 March 1946]

1. The general non-linear difference-differential equation with constant coefficients may be written in the form

$$\Lambda_1\{y(x)\}+\Lambda_2\{y(x)\} = v(x), \tag{1.1}$$

where
$$\Lambda_1\{y(x)\} \equiv \sum_{\mu=0}^{m} \sum_{\nu=0}^{n} a_{\mu\nu} y^{(\nu)}(x+b_\mu),$$
$$\Lambda_2\{y(x)\} \equiv \sum_{\lambda} A_\lambda y^{(\beta_{\lambda,1})}(x+b_{\lambda,1}) y^{(\beta_{\lambda,2})}(x+b_{\lambda,2})\dots,$$
$$y^{(0)}(x) \equiv y(x),$$

and $v(x)$ is a known function of the real variable x. We suppose that $m \geqslant 1$, $n \geqslant 1$ and $0 \leqslant \beta_{\lambda,i} \leqslant n$. The numbers a, b, A are independent of x, and the b are real. $\Lambda_2\{y(x)\}$ has a finite number of terms and each term has at least two y-function factors.

Particular examples of (1.1) have occurred in arithmetic, in the problem of the iterated exponential function and the associated Schroeder's equation, and in economics. The theory of the general equation is complicated, as we might expect, but I have developed a fairly straightforward theory of those solutions which tend to a finite limit as $x \to \infty$. An obvious transformation reduces this to the case in which the limit is zero and the solution is small as $x \to \infty$. Our problem has three stages.

In the first we have to determine, for any particular example of (1.1), whether there is a solution which is small as $x \to \infty$ and what conditions in some finite interval of x will ensure that the solution is of this type. Apart from trivialities, this appears to be a genuinely difficult problem and one which we can only hope to solve in particular cases. When this problem can be solved for a particular equation, however, we may be able to obtain, by means of the two subsequent stages, a complete solution of the equation, that is, an explicit formulation of all the solutions. I hope to illustrate this idea elsewhere by a discussion of the equation

$$y'(x+1) = -\alpha y(x)\{1+y(x+1)\},$$

which I have already mentioned.*

* Wright, 'On a sequence defined by a non-linear recurrence formula', *J. of London Math. Soc.* 20 (1945), 68–73.

The second problem is to show that, under suitable general conditions, any small solution must be exponentially small, that is $O(e^{-cx})$ for some $c > 0$ as $x \to \infty$. The third is to find an asymptotic expansion for all exponentially small solutions. I give here a fairly general solution of the third problem and postpone consideration of the second problem for the present.*

We considerably simplify the statement of the result and the proof by confining ourselves to the equation

$$\Lambda_3\{y(x)\} + \Lambda_2\{y(x)\} = v(x), \tag{1.2}$$

where
$$\Lambda_3\{y(x)\} \equiv y^{(n)}(x) + \sum_{\mu=1}^{m} \sum_{\nu=0}^{n-1} a_{\mu\nu}\, y^{(\nu)}(x+b_\mu)$$

and $\Lambda_2\{y(x)\}$ is unaltered. The treatment of (1.2) exhibits the essentials of my method without being unduly lengthy. For this equation I show that, if $0 < c < C$, $y^{(\nu)}(x) = O(e^{-cx})$ for $0 \leqslant \nu \leqslant n$ and $v(x) = O(e^{-Cx})$ as $x \to \infty$, then we can find asymptotic expansions for the $y^{(\nu)}(x)$, the error being $O(e^{-(C-\epsilon)x})$ for any $\epsilon > 0$.

This result can be extended, at the cost of an increased complexity of statement and proof, to more general equations, for example (1.1) and those in which the a and the A are suitable known functions of x.

2. The nature of the result will be clearer if we examine briefly the *associated linear equation*†

$$\Lambda_3\{y(x)\} = 0 \tag{2.1}$$

and the *associated transcendental equation*

$$\tau(s) \equiv s^n + \sum_{\mu=1}^{m} \sum_{\nu=0}^{n-1} a_{\mu\nu}\, s^\nu e^{b_\mu s} = 0. \tag{2.2}$$

Clearly
$$\Lambda_3(e^{sx}) = \tau(s)e^{sx}.$$

Also, if M is any positive integer,

$$\Lambda_3(x^M e^{sx}) = \Lambda_3\left(\frac{\partial^M}{\partial s^M} e^{sx}\right) = \frac{\partial^M}{\partial s^M} \Lambda_3(e^{sx}) = e^{sx} \sum_{\lambda=0}^{M} \binom{M}{\lambda} \tau^{(M-\lambda)}(s)x^\lambda. \tag{2.3}$$

* When I first submitted this article to the editors, I could only solve the second problem in a special case; I have since found a general solution.

† This equation and the equation $\Lambda_1\{y(x)\} = 0$ have been discussed by many authors, especially Schmidt, *Math. Ann.* 70 (1911), 499–524; Hilb, ibid. 78 (1918), 137–70; Titchmarsh, *Fourier Integrals* (Oxford, 1937) and *J. of London Math. Soc.* 14, 118–24; and Pitt, *Proc. Cambridge Phil. Soc.* 40 (1944), 199–211, and in a paper as yet unpublished.

Hence, if s_r is any zero of $\tau(s)$, p_r+1 its order and $P_r(x)$ any polynomial of degree p_r (or less) in x, $P_r(x)e^{s_r x}$ is a solution of (2.1). So is

$$\sum_{s_r} P_r(x)e^{s_r x}, \tag{2.4}$$

the sum containing a finite or, under suitable conditions as to convergence, an infinite number of terms.

3. Let $0 < c < C$. We shall see (Lemma 1) that the number of zeros s_r of $\tau(s)$, whose real parts lie in any finite interval, is finite.* Hence we may denote by

$$s_1,\ s_2,\ \ldots,\ s_R, \tag{3.1}$$

where $\mathscr{R}(s_{r+1}) \leqslant \mathscr{R}(s_r)$, the zeros of $\tau(s)$ for which

$$-C < \mathscr{R}(s) \leqslant -c. \tag{3.2}$$

We now construct the set of numbers S, each member of which is formed by adding together two or more of the numbers (3.1), repetitions being allowed, so that $2s_1$, s_1+s_2, $3s_1$, for example, are members. Only a finite number of the S will satisfy (3.2) and these we denote by

$$S_1,\ S_2,\ \ldots,\ S_{R'}, \tag{3.3}$$

where $\mathscr{R}(S_{r+1}) \leqslant \mathscr{R}(S_r)$.

In what follows x and T are real variables and $s = \sigma+it$ is a complex variable; ν is a whole number ($0 \leqslant \nu \leqslant n$), and any statement involving ν is true for all such ν unless the contrary is stated. We use ϵ to denote a positive number (to be thought of as small) and K (not always the same at each occurrence) to denote a positive number, independent of x, s, T, but depending on all other parameters present, including ϵ. The notation $O(\quad)$ refers to the passage of x to ∞, and the constant implied is of the type K. All statements such as $f(x) = O(e^{\epsilon x})$ are to be read as true for every fixed $\epsilon > 0$. The numbers $K_1, K_2, \ldots$ are of the type K, but K_1, for example, has the same value at each occurrence.

The main result of this paper is the theorem:

THEOREM. *If* (i) $y(x)$ *is a solution of* (1.2), (ii) $v(x)$ *and* $y^{(\nu)}(x)$ *are continuous for* $x > K_1$ *and of bounded variation in any finite interval* (K_1, X), *where* $X > K_1$, *and* (iii) $v(x) = O(e^{-(C-\epsilon)x})$, $y^{(\nu)}(x) = O(e^{-(c-\epsilon)x})$, *then there are polynomials*

$$P_r(x) \quad (1 \leqslant r \leqslant R), \qquad Q_r(x) \quad (1 \leqslant r \leqslant R')$$

* This is not true for the associated transcendental equation of (1.1).

and a sum $$\eta(x) = \sum_{r=1}^{R} P_r(x)e^{s_r x} + \sum_{r=1}^{R'} Q_r(x)e^{S_r x}$$

such that $$y^{(\nu)}(x) = \eta^{(\nu)}(x) + O(e^{-(C-\epsilon)x}).$$

The degree of the polynomial $P_r(x)$ is at least one less than the order of the zero $s = s_r$ of $\tau(s)$.

If we have no information about $y(x)$ beyond that given in the theorem, the coefficients in the $P_r(x)$ are undetermined. But, if these coefficients are assigned, the degrees and coefficients of the $Q_r(x)$ can be determined by equating coefficients in the relation

$$\Lambda_3\{\eta(x)\} + \Lambda_2\{\eta(x)\} = O(e^{-(C-\epsilon)x}).$$

There is no loss of generality if we replace K_1 in hypothesis (ii) by $b = \min(0, b_1, \dots, b_m)$, since this involves only a change of origin of x.

4. LEMMA 1. *The function $\tau(s)$ has only a finite number of zeros for which $|\sigma| \leqslant K_2$. Corresponding to any K_2 there exist numbers K_3 and K_4 such that $|\tau(s)| > K_3|s^n|$ whenever $|\sigma| \leqslant K_2$, $|t| > K_4$.*

When $|\sigma| \leqslant K_2$, we have

$$|\tau(s) - s^n| < K(1 + |s^{n-1}|).$$

The second part of the lemma is immediate. Hence all the zeros of $\tau(s)$ for which $|\sigma| \leqslant K_2$ must lie in the rectangle $|\sigma| \leqslant K_2$, $|t| \leqslant K_4$. Since $\tau(s)$ is regular, these zeros must be finite in number.

LEMMA 2. *Let L be any non-negative integer, and S be any number (real or complex). If $\tau(S) = 0$, let ω be the order of the zero $s = S$ of $\tau(s)$, while, if $\tau(S) \neq 0$, let $\omega = 0$. We can determine in succession numbers $B_L, B_{L-1}, \dots, B_0$ such that**

$$\Lambda_3\left\{e^{Sx} \sum_{l=0}^{L} \frac{B_l}{(\omega+l)!} x^{l+\omega}\right\} = x^L e^{Sx}.$$

We observe that $\tau^{(\omega)}(S) \neq 0$. Let

$$\rho_l = \frac{\tau^{(l+\omega)}(S)}{(l+\omega)!},$$

so that $\rho_0 \neq 0$, and choose $B_L, B_{L-1}, \dots, B_0$ so that

$$B_L \rho_0 = L!, \qquad B_{L-v}\rho_0 = -\sum_{u=0}^{v-1} B_{L-u}\rho_{v-u} \quad (1 \leqslant v \leqslant L). \tag{4.1}$$

* As usual, $0! = 1$.

By (2.3)

$$e^{-Sx}\Lambda_3\left\{e^{Sx}\sum_{l=0}^{L}\frac{B_l}{(\omega+l)!}x^{l+\omega}\right\} = \sum_{l=0}^{L} B_l \sum_{\lambda=0}^{l} \rho_{l-\lambda}\frac{x^\lambda}{\lambda!}$$
$$= \sum_{\lambda=0}^{L}\frac{x^\lambda}{\lambda!}\sum_{l=\lambda}^{L} B_l\rho_{l-\lambda}$$
$$= \sum_{v=0}^{L}\frac{x^{L-v}}{(L-v)!}\sum_{u=0}^{v} B_{L-u}\rho_{v-u},$$

when we put $\lambda = L-v$, $l = L-u$. By (4.1) this reduces to x^L.

5. Lemma 3. *If* (i) $d_1 > d_2$, (ii) $w(x)$ *and* $z^{(\nu)}(x)$ *are continuous and of bounded variation in any finite interval* (b, X) *where* $X > b$,

(iii) $$\Lambda_3\{z(x)\} = w(x) \quad (x \geqslant b), \tag{5.1}$$

(iv) $$w(x) = O(e^{(d_2+\epsilon)x}), \qquad z^{(\nu)}(x) = O(e^{(d_1+\epsilon)x}),$$

(v) $$s_{j+1},\ldots, s_J$$

are the zeros of $\tau(s)$ *for which* $d_2 < \sigma \leqslant d_1$, *then there are polynomials* $P_r(x)$ $(j+1 \leqslant r \leqslant J)$, *of the same degree as before, such that, if*

$$\zeta(x) = \sum_{r=j+1}^{J} P_r(x)e^{s_r x},$$

then $$z^{(\nu)}(x) = \zeta^{(\nu)}(x)+O(e^{(d_2+\epsilon)x}). \tag{5.2}$$

We can choose ϵ small enough to ensure that $\tau(s)$ has no zeros for which

$$d_1 < \sigma \leqslant \sigma_1 = d_1+2\epsilon, \qquad d_2 < \sigma \leqslant \sigma_2 = d_2+2\epsilon,$$

and that $\sigma_2 \neq 0$. We can suppose without loss of generality that

$$z(0) = z'(0) = \ldots = z^{(n-1)}(0) = 0, \tag{5.3}$$

since, if not, we have only to replace $z(x)$ by

$$z(x)-\pi(x)e^{d_1 x},$$

where $\pi(x)$ is a suitable polynomial. This affects neither the hypotheses nor the conclusions of Lemma 3.

Let

$$H(s) = \sum_{\mu=1}^{m}\sum_{\nu=0}^{n-1} a_{\mu\nu}e^{b_\mu s}\int_0^{b_\mu} z^{(\nu)}(x)e^{-sx}\,dx,$$

$$Z_\nu(s) = \int_0^\infty z^{(\nu)}(x)e^{-sx}\,dx \quad (\mathscr{R}(s) \geqslant \sigma_1),$$

$$W(s) = \int_0^\infty w(x)e^{-sx}\,dx \quad (\mathscr{R}(s) \geqslant \sigma_2).$$

Integrating by parts and using (5.3), we see that

$$Z_\nu(s) = s^\nu Z(s). \tag{5.4}$$

Multiplying (5.1) by e^{-sx} and integrating with respect to x from 0 to ∞ we have

$$\tau(s)Z(s) = H(s)+W(s). \tag{5.5}$$

LEMMA 4. *$H(s)$ is an integral function of s and, for bounded σ, $sH(s)$ is bounded.*

The first part is obvious. For the second, we have

$$sH(s) = \sum_{\mu=1}^{m}\sum_{\nu=0}^{n-1} a_{\mu\nu}\left\{e^{b_\mu s}\int_0^{b_\mu} z^{(\nu+1)}(x)e^{-sx}\,dx - z^{(\nu)}(b_\mu)\right\},$$

on integrating by parts. By hypothesis (ii) of Lemma 3 this is bounded.

LEMMA 5. *If $\sigma \geqslant \sigma_2$, then $W(s)$ is a regular function of s and tends to zero uniformly in σ as $|t| \to \infty$. Also*

$$\int_{-\infty}^{\infty}\left|\frac{W(\sigma_2+it)}{\sigma_2+it}\right| dt$$

is convergent.

These results all follow by standard theorems from hypotheses (ii) and (iv) of Lemma 3 and the definition of $W(s)$.

We take $K_2 > \max(|\sigma_1|, |\sigma_2|)$ in Lemma 1, $T > K_4$ and $P_r(x)e^{s_r x}$ to be the residue of

$$\left\{\frac{H(s)+W(s)}{\tau(s)}\right\}e^{sx} = Z(s)e^{sx}$$

at $s = s_r$. By Cauchy's theorem, the integral of $Z(s)s^\nu e^{sx}$ round the rectangle whose vertices are $\sigma_1 \pm iT$, $\sigma_2 \pm iT$ is equal to $2\pi i\zeta^{(\nu)}(x)$.

Now let $\nu \leqslant n-1$. By Lemmas 1, 4 and 5,

$$\int_{\sigma_2-iT}^{\sigma_1-iT} + \int_{\sigma_1+iT}^{\sigma_2+iT} Z(s)s^\nu e^{sx}\,ds \to 0$$

as $T \to \infty$, and

$$\left|\int_{\sigma_2-iT}^{\sigma_2+iT} Z(s)s^\nu e^{sx}\,ds\right| \leqslant Ke^{\sigma_2 x}\int_{-T}^{T}\left|\frac{H(\sigma_2+it)+W(\sigma_2+it)}{\sigma_2+it}\right| dt$$

$$\leqslant Ke^{\sigma_2 x} = Ke^{(d_2+2\epsilon)x}.$$

Hence

$$2\pi i\zeta^{(\nu)}(x) = \lim_{T\to\infty}\int_{\sigma_1-iT}^{\sigma_1+iT} Z(s)s^\nu e^{sx}\,ds + O(e^{(d_2+2\epsilon)x})$$
$$= 2\pi i z^{(\nu)}(x)+O(e^{(d_2+2\epsilon)x}),$$

by the usual Laplace inversion-formula. Since this is true for every $\epsilon > 0$, we have (5.2) for $\nu \leqslant n-1$.

By (2.4) and (5.1), we have

$$\Lambda_3\{\zeta(x)\} = 0, \qquad \Lambda_3\{z(x)-\zeta(x)\} = w(x) = O(e^{(d_2+\epsilon)x}).$$

Since (5.2) is true for $\nu \leqslant n-1$, it follows that

$$z^{(n)}(x)-\zeta^{(n)}(x) = O(e^{(d_2+\epsilon)x}),$$

which is (5.2) for $\nu = n$.

6. Let k be any positive integer such that $kc < C$ and let

$$D = \min\{C, (k+1)c\}.$$

Let
$$s_1, \ldots, s_j; \qquad S_1, \ldots, S_{j'} \tag{6.1}$$
be the s_r and S_r for which* $-kc < \sigma \leqslant -c$ and let

$$s_{j+1}, \ldots, s_J; \qquad S_{j'+1}, \ldots, S_{J'}$$

be the s_r and S_r for which $-D < \sigma \leqslant -kc$. We write

$$\eta_1(x) = \sum_{r=1}^{j} P_r(x)e^{s_r x}+\sum_{r=1}^{j'} Q_r(x)e^{S_r x},$$

and $\eta_2(x)$ for the same sum with J, J' replacing j, j'.

LEMMA 6. *If the hypotheses of our theorem are satisfied and if, in addition,*

$$y^{(\nu)}(x) = \eta_1^{(\nu)}(x)+O(e^{-(kc-\epsilon)x}),$$

then
$$y^{(\nu)}(x) = \eta_2^{(\nu)}(x)+O(e^{-(D-\epsilon)x}).$$

Let us write
$$y(x) = \eta_1(x)+y_1(x),$$
so that
$$y_1^{(\nu)}(x) = O(e^{-(kc-\epsilon)x}).$$
Also $\eta_1^{(\nu)}(x) = O(e^{-(c-\epsilon)x})$. Now

$$\Lambda_2(y)-\Lambda_2(\eta_1) = \Lambda_2(\eta_1+y_1)-\Lambda(\eta_1)$$

is the sum of a finite number of terms, each the product of two or more factors, one or more factors being $y_1^{(\nu)}$ functions and the remaining factors $\eta_1^{(\nu)}$ functions. Such a term is

$$O(e^{-\{(k+1)c-2\epsilon\}x}) = O(e^{-(D-\epsilon)x}),$$

* If $k = 1$, this is to be read as making (6.1) empty.

since ϵ is *any* positive number. Hence, by (1.2),

$$\Lambda_3(y_1) = -\Lambda_3(\eta_1)-\Lambda_2(\eta_1)-\{\Lambda_2(y)-\Lambda_2(\eta_1)\}+v(x)$$
$$= \sum_{r=1}^{j} P_{r,1}(x)e^{s_r x}+\sum_{r=1}^{J'} Q_{r,1}(x)e^{S_r x}+O(e^{-(D-\epsilon)x})$$

for some polynomials $P_{r,1}$, $Q_{r,1}$. Now

$$\Lambda_3\{y_1\} = O(e^{-(kc-\epsilon)x})$$

and so $$\sum_{r=1}^{j} P_{r,1}(x)e^{s_r x}+\sum_{r=1}^{j'} Q_{r,1}(x)e^{S_r x} = O(e^{-(kc-\epsilon)x}) \tag{6.2}$$

for any $\epsilon > 0$. Since the real part of each of the numbers of (6.1) is greater than $-kc$, the left-hand side of (6.2) must be identically zero. We have therefore

$$\Lambda_3\{y_1(x)\} = \sum_{r=j'+1}^{J'} Q_{r,1}(x)e^{S_r x}+O(e^{-(D-\epsilon)x}).$$

By Lemma 2, we can find polynomials $Q_r(x)$ ($j' < r \leqslant J'$) such that

$$\Lambda_3\Big\{\sum_{r=j'+1}^{J'} Q_r(x)e^{S_r x}\Big\} = \sum_{r=j'+1}^{J'} Q_{r,1}(x)e^{S_r x}.$$

Putting $$y_1(x) = \sum_{r=j'+1}^{J'} Q_r(x)e^{S_r x}+z(x),$$

we have $$\Lambda_3\{z(x)\} = O(e^{-(D-\epsilon)x})$$

and $$z^{(\nu)}(x) = O(e^{-(kc-\epsilon)x}).$$

Since $y^{(\nu)}(x)$ and $P(x)e^{sx}$ are continuous for $x \geqslant b$ and of bounded variation in any finite interval (b, X), the same is true for $z^{(\nu)}(x)$. We now put $d_1 = -kc$ and $d_2 = -D$ in Lemma 3. The hypotheses of that lemma are satisfied and so

$$z^{(\nu)}(x) = \frac{d^\nu}{dx^\nu}\sum_{r=j+1}^{J} P_r(x)e^{s_r x}+O(e^{-(D-\epsilon)x}).$$

This completes the proof of Lemma 6.

To prove our theorem we have only to apply Lemma 6 repeatedly with $k = 1, 2,..., k_0$, where $(k_0+1)c \geqslant C$.

ON FROBENIUS' THEOREM

By C. S. FU (*Peiping*)

[Received 1 March 1946]

In a group $\mathfrak{G}$ of order ab let N_b be the number of elements of orders dividing b. The fundamental theorem that 'N_b is divisible by b' was discovered by Frobenius* in 1895. We shall be concerned with the question '*when is N_b equal to b?*' One of the answers, due to Frobenius,† is that '*if a has no equal factors and every prime factor of a is less than every factor* (> 1) *of b, then $N_b = b$*'. This theorem was generalized by me in 1941 into the form:

THEOREM 1. *Let $\mathfrak{G}$ be a group of order ab and N_b be the number of elements in $\mathfrak{G}$ whose orders divide b. If b is prime to $a\phi(a)$, and no two distinct cyclic sub-groups conjugate in $\mathfrak{G}$ occur in the same Sylow's sub-group of order dividing a, then $N_b = b$.*

[1] This is evidently true when $a = 1$.

I shall complete induction by proving

[2] *The theorem is true for a if it is true for all smaller values of a.*

[3] Let p be the greatest prime factor of a, p^e the highest power of p dividing a, and $\mathfrak{H}$ a Sylow's sub-group of order p^e. We can easily find distinct elements in $\mathfrak{H}$

$$P_1, \quad ..., \quad P_s \quad (P_1 = E) \tag{1}$$

such that the corresponding groups generated separately by them

$$\{P_1\}, \quad ..., \quad \{P_s\} \tag{2}$$

form the total set of distinct cyclic sub-groups of $\mathfrak{H}$. Let p^{e_i} be the order of P_i. Then $\{P_i\}$ contains exactly $\phi(p^{e_i})$ elements of order p^{e_i},

$$P_i^j, \tag{3}$$

where j runs over all positive integers less than and prime to p^{e_i}. Thus $\{P\} = \{P_i\}$ if and only if P is one of the elements (3). Hence

$$P_i^j \quad \begin{pmatrix} i = 1, ..., s, \\ j < p^{e_i} \text{ and prime to it} \end{pmatrix} \tag{4}$$

are the p^e distinct elements of $\mathfrak{H}$ and

$$p^e = \sum_{i=1}^{s} \phi(p^{e_i}). \tag{5}$$

[4] By hypothesis no two of the sub-groups (2) are conjugate in $\mathfrak{G}$. Hence each $\{P_i\}$ is invariant in $\mathfrak{H}$, i.e. $\mathfrak{H}$ is contained in the normalizer

* Cf. Burnside, *Theory of Groups*, § 87, Theorem VII.

† Cf. M. Sono, *Theory of Groups*.

$\mathfrak{N}_i$ of $\{P_i\}$ in $\mathfrak{G}$. The order and the index of $\mathfrak{N}_i$ may therefore be written respectively in the forms $a_i p^e b_i$ and $\alpha_i \beta_i$, in which $a = a_i p^e \alpha_i$, $b = b_i \beta_i$.

[5] According to the first theorem of Frobenius, quoted above, $\mathfrak{N}_i$ contains $m_i b_i$ distinct elements

$$Q_k \quad (k = 1,..., m_i b_i) \tag{6}$$

such that $Q_k^{b_i} = E$ (which is equivalent to $Q_k^b = E$).

[6] Again, *$p^e b$ is divisible by the orders of the elements*

$$P_i^j Q_k \quad \begin{pmatrix} j < p^{e_i} \text{ and prime to it,} \\ k = 1,..., m_i b_i. \end{pmatrix} \tag{7}$$

For, $Q_k^{-1}\{P_i\}Q_k = \{P_i\}$, we may set $Q_k^{-1} P_i Q_k = P_i^t$ ($t < p^{e_i}$ and prime to it). Then

$$P_i = Q_k^{-b} P_i Q_k^b = P_i^{t^b}, \qquad 1 \equiv t^b \pmod{p^{e_i}}.$$

Hence either (i) $t = 1$ or

(ii) $\quad t \neq 1, \qquad t^b \equiv 1 \pmod{p^{e_i}}, \qquad b \equiv 0 \pmod{\phi(p^{e_i})}$.

The case (ii) is impossible since b is prime to $a\phi(a)$ by hypothesis. Thus $t = 1$,

$$P_i Q_k = Q_k P_i, \qquad (P_i^j Q_k)^{p^e b} = P_i^{jp^e b} Q_k^{p^e b} = E.$$

[7] If $R_1,..., R_{\alpha_i \beta_i}$ are elements of $\mathfrak{G}$ such that

$$R_l^{-1} \mathfrak{N}_i R_l \quad (l = 1,..., \alpha_i \beta_i) \tag{8}$$

form the complete set of conjugates in $\mathfrak{G}$ to which $\mathfrak{N}_i$ belongs, then

$$R_l^{-1}\{P_i\}R_l \quad (l = 1,..., \alpha_i \beta_i) \tag{9}$$

form the complete set of conjugates in $\mathfrak{G}$ to which $\{P_i\}$ belongs. Further, by [6], *$p^e b$ is divisible by the orders of the*

$$\sum_{i=1}^{s} \phi(p^{e_i}) m_i b_i \alpha_i \beta_i = \sum_{i=1}^{s} \phi(p^{e_i}) m_i \alpha_i b \tag{10}$$

elements

$$R_l^{-1} P_i^j Q_k R_l \quad \begin{pmatrix} i = 1,..., s, \\ j < p^{e_i} \text{ and prime to it,} \\ k = 1,..., m_i b_i, \\ l = 1,..., \alpha_i \beta_i. \end{pmatrix} \tag{11}$$

[8] *Every element C whose order divides $p^e b$ is contained in* (11).

Since, by hypothesis, p^e is prime to b, there exist integers x and y such that $p^e x + by = 1$. Put

$$A = C^{by}, \qquad B = C^{p^e x};$$

then $$A^{p^e} = E = B^b, \qquad AB = C = BA.$$
Hence $\{A\}$ is conjugate to one of the sub-groups (2). Let
$$\{A\} = R_l^{-1}\{P_i\}R_l, \qquad A = R_l^{-1}P_i^j R_l, \qquad B = R_l^{-1}QR_l;$$
then $$Q^b = E, \qquad P_i^j Q = QP_i^j, \qquad P_i Q = QP_i.$$
Hence Q is one of the elements (6). Let $Q = Q_k$, then
$$C = AB = R_l^{-1}P_i^j Q_k R_l.$$

[9] *The elements* (11) *are distinct.*

For, if two elements in (11), $C = R_l^{-1}P_i^j Q_k R_l$ and $R_{l'}'^{-1}P_{i'}^{j'} Q_{k'}' R_{l'}'$, are equal; then the elements
$$A = R_l^{-1}P_i^j R_l, \qquad B = R_l^{-1}Q_k R_l,$$
$$A' = R_{l'}'^{-1}P_{i'}^{j'} R_{l'}', \qquad B' = R_{l'}'^{-1}Q_{k'}' R_{l'}'$$
have the relations
$$A^p = B^b = E = A'^{p^e} = B'^b, \qquad AB = BA = C = A'B' = B'A'.$$
If $p^e x + by = 1$, then $A = C^{by} = A'$, $B = C^{p^e x} = B'$. Further, from $A = A'$, we have
$$R_l^{-1}\{P_i^j\}R_l = R_{l'}'^{-1}\{P_{i'}^{j'}\}R_{l'}'.$$
But, by hypothesis, $\mathfrak{H}$ contains no distinct cyclic sub-groups conjugate in $\mathfrak{G}$. Hence $\{P_i^j\} = \{P_{i'}^{j'}\}$, $\{P_i\} = \{P_{i'}\}$, $P_i = P_{i'}$, by [3]. Again, from $R_l^{-1}\{P_i\}R_l = R_{l'}'^{-1}\{P_{i'}\}R_{l'}'$, we have $R_l = R_{l'}'$, by [7]. Consequently $P_i^j = P_{i'}^{j'}$ (from $A = A'$) and $Q_k = Q_{k'}'$ (from $B = B'$).

[10] By [7], [8], and [9] the number of elements in $\mathfrak{G}$ whose orders divide $b' = p^e b$ is
$$N_{b'} = \sum_{i=1}^{s} \phi(p^{e_i}) m_i \alpha_i b. \tag{12}$$

[11] Put $a' = a/p^e$, $b' = p^e b$. Then $a'b'$ is the order of $\mathfrak{G}$. It is evident that two distinct cyclic sub-groups in the same Sylow's sub-group of order dividing a' are not conjugate in $\mathfrak{G}$. Further, since p is greater than every prime factor of a', the integer b' must be prime to $a'\phi(a')$. Hence, by the inductive hypothesis of [2], $N_{b'} = b' = p^e b$. Consequently, from (12) and (5),
$$\sum_{i=1}^{s} \phi(p^{e_i}) m_i \alpha_i = p^e = \sum_{i=1}^{s} \phi(p^{e_i}).$$
This gives each $m_i \alpha_i = 1$, $m_i = 1$, $\alpha_i = 1$.

[12] By [11], $m_1 = 1$. By [4], $\mathfrak{N}_1 = \mathfrak{G}$, $b_1 = b$. Hence, by [5], $b = b_1 m_1$ is the number of elements Q in $\mathfrak{G}$ such that $Q^b = E$. This proves $b = N_b$.

As particular cases of Theorem 1 I mention three corollaries:

Corollary 1. *If b is prime to $a\phi(a)$ and every Sylow's sub-group of order dividing a is cyclic,* then $N_b = b$.*

Corollary 2. *If a has no equal factors and b is prime to $a\phi(a)$, then $N_b = b$.*

Corollary 3. *If a has no equal factors and every prime factor of a is less than every factor (> 1) of b, then $N_b = b$.*

Corollary 3 is the second theorem of Frobenius, quoted above, and is a particular case of Corollary 2. As an example, Corollary 3 cannot be applied to the case

$$ab = 90 = 2.5.3^2.$$

But, according to Corollary 2, for every group $\mathfrak{G}$ of order $90 = 2.5.3^2$ we have $N_9 = 9$, $N_{45} = 45$. Further, if $\mathfrak{G}_1$ and $\mathfrak{G}_2$ are the Sylow's sub-groups of orders 9 and 5 respectively; then $\mathfrak{G}_1$, and therefore $\mathfrak{G}_1\mathfrak{G}_2 = \mathfrak{G}_2\mathfrak{G}_1$, must be invariant in $\mathfrak{G}$. Again, since

$$1+5 < 9 < 1+2.5 < 2(1+5),$$

$\mathfrak{G}_2$ is invariant in $\mathfrak{G}_1\mathfrak{G}_2$ by Sylow's theorem. Hence $\mathfrak{G}_2$ is also invariant in $\mathfrak{G}$. (This was not recognized by M. Sono in his *Theory of Groups*, § 59 and § 78.) Similar conclusions may be obtained for the case $ab = 2.17.5.3^2 = 2.5.17.3^2$.

In Theorem 1 put $ab = cq_1^{r_1}\dots q_m^{r_m}$ and $b = q_m^{r_m}$, $q_{m-1}^{r_{m-1}}q_m^{r_m}, \dots$ successively, and we have

Theorem 2. *If $q_1, \dots, q_m$ are distinct primes not dividing $c\phi(c)$ and each q_μ $(1 < \mu \leqslant m)$ does not divide $(q_1-1)\dots(q_{\mu-1}-1)$, and if $\mathfrak{G}$ is a group of order $cq_1^{r_1}\dots q_m^{r_m}$ in which no two conjugate cyclic sub-groups are contained in the same Sylow's sub-group of order dividing $cq_1^{r_1}\dots q_{m-1}^{r_{m-1}}$; then $\mathfrak{G}$ contains an invariant sub-group of order*

$$q_\lambda^{r_\lambda}\dots q_m^{r_m} \quad (\lambda = m, m-1, \dots, 1)$$

involving all the elements of $\mathfrak{G}$ whose orders divide that order.

This theorem has also corollaries similar to those above. For example, the theorem given by Hilton, quoted in the footnote, is a very particular case of Theorem 2.

* One of the essential distinctions between this and the theorem given by Hilton in his *Finite Groups*, chap. xiv, § 4, is that I do not assume the Sylow's sub-groups of orders dividing b to be cyclic.

W-

C51